MAY 2 1921
MAY 2 1921
UNIV. OF MICH.

Journal of Educational Research

Official Organ of the National Association of Directors of Educational Research

B. R. BUCKINGHAM, *Editor*

Associate Editors

E. J. Ashbaugh — W. W. Charters — George D. Strayer
Leonard P. Ayres — S. A. Courtis — Lewis M. Terman
Walter S. Monroe

VOL. III APRIL, 1921 No. 4

CONTENTS

RECENT DEVELOPMENTS IN MEASURING HUMAN CAPACITIES ..... M. E. Haggerty 241

THE EFFECT OF CLEAR OBJECTIVES ON THE TEACHING OF READING ..... William E. Hawley 254

AN EXPERIMENT WITH THE OTIS GROUP INTELLIGENCE SCALE IN THE NEEDHAM, MASSACHUSETTS, HIGH SCHOOL ..... Roscoe L. West 261

RAYMOND: A CASE STUDY ILLUSTRATIVE OF THE DIAGNOSTIC LIMITATIONS OF A QUANTITATIVE SCORE ..... Bernice Leland 280

RATING THE GENERAL APPEARANCE OF PRIMARY NUMBER WORK ..... Ernest C. Witham 273

METHODS OF INTERPRETING RESULTS IN THE CLEVELAND ARITHMETIC TESTS ..... Joseph Peterson 280

EDITORIALS 293

REVIEWS AND ABSTRACTS 300

NEWS ITEMS AND COMMUNICATIONS 305

NATIONAL ASSOCIATION OF DIRECTORS OF EDUCATIONAL RESEARCH 317

AMERICAN BOOK COMPANY

EDUCATIONAL PUBLISHERS

NEW YORK, 100 Washington Square
CINCINNATI, 300 Pike Street
CHICAGO, 330 East 22nd Street
BOSTON, 63 Summer Street
ATLANTA, 2 North Forsyth Street

Address all communications to the nearest office of the company. Descriptive circulars, catalogues, and price lists will be mailed on request to any address.

Your correspondence is solicited and will have prompt attention.

J. B. LIPPINCOTT COMPANY

Wishes to call your special attention to the following publications for library use:

BRIGHTNESS AND DULLNESS IN CHILDREN.....$1.40
By Herbert Woodrow, Associate Professor of Psychology, University of Minnesota.

THINKING AND LEARNING TO THINK.....$1.40
By N. C. Schaeffer, Ph.D., LL.D., Superintendent of Public Instruction for the State of Pennsylvania.

SOCIOLOGICAL DETERMINATION OF OBJECTIVES IN EDUCATION.....$2.50
By David Snedden, Professor of Education, Columbia University.

REDIRECTION OF HIGH SCHOOL INSTRUCTION.....$1.60
By Lull (Kansas State Normal School), and Wilson (Superintendent of Schools, Berkeley, Cal.)

TRAINING FOR A LIBRARIANSHIP.....$1.75
By Friedel.

Write for further information.

J. B. LIPPINCOTT COMPANY
2126 Prairie Avenue CHICAGO, ILLINOIS

JOURNAL *of* EDUCATIONAL RESEARCH

VOLUME III APRIL, 1921 NUMBER 4

RECENT DEVELOPMENTS IN MEASURING HUMAN CAPACITIES[1]

M. E. HAGGERTY

University of Minnesota

This association requires its president to give an address. From the form of the resolution by which this custom was inaugurated it was obviously the original intent of the association that the address should be a review of educational progress in the calendar year preceding the annual meeting. That such a review can be comprehensive, detailed, and critical is of course impossible. Not only the volume of recent educational literature but its diversity of content and its range of method precludes the accomplishment of a purpose so elaborate. The allotted time requires a more restricted project. With your permission, I propose, therefore, to select a limited portion of the field which might legitimately be included in such a review and to confine myself to a discussion of recent events in the attempt to measure human capacities. I shall endeavor to indicate something of the progress made during the past year, to point out certain of the newer interests, and to evaluate our present status. If a single summary phrase were useful to indicate the drift of current discussion we might choose "the Inadequacy of Intelligence" as a suitable title for this paper.

Let us first give attention to the extension of intelligence examinations which has proceeded with acceleration during the year. The avidity with which the educational public has seized upon group intelligence examinations is both encouraging and alarming. It confirms our faith that such tests meet a real need in school work, but it also raises a doubt as to the existence of a wholesome critical attitude of mind toward the proper selecting of tests and the proper use of test results.

[1] Address of the retiring president of the National Association of Directors of Educational Research, delivered at the annual meeting of the association, at Atlantic City, N. J., March 3, 1921.

So widespread is the interest in the measurement of intelligence of young people in the schools that we may speak of it as a current educational movement. Students of education have devised the means—the tests, scales, and examinations—and practical educators are using them in the belief that a proper evaluation of the intelligence of pupils will make possible important remedial measures in teaching and school organization. It may be worth while briefly to examine the chief characteristics of this movement.

In the first place it is obvious that interest in the measurement of intelligence has been greatly stimulated by events connected with the great war. For a decade prior to that event there had been growing in the minds of students of education a conviction that the results of formal school education were greatly influenced by the native capacities of children. This conviction was largely due to the growth of intelligence tests and the revelations which such tests gave of the wide variabilities in mental capacity to be found among the school population. This background of belief was fertile soil upon which fell the spectacular examination of 1,700,000 soldiers by group intelligence examinations during the war. Educators were quick to see the possibilities of such tests as means of improving educational processes. Teachers, administrators, and supervisors, as well as students of psychology and education, have received the adaptations of the group intelligence examination to school uses with open arms and all too often with uncritical acceptance of what has been made available.

The eagerness of the educational public for these new scales and tests is evident from the number sold. When the first group scale was offered for commercial distribution the publishers accepted the offer as an adventure, printing a first lot of 10,000. Contrast this cautious move with the sales of the National Intelligence Tests first put on the market in September, 1920. More than 200,000 each of the first forms of scales A and B have already been distributed. This large sale occurs notwithstanding the rapidly increasing number of other scales offered in competition. It is probable that well-nigh 4,000,000 pupils in the public schools of the United States will have been examined by reliable group intelligence examinations by the end of the present academic year, less than 30 months after the publication of the first scale of this type.

A second fact to be noted is the range of individuals for whom standard group intelligence tests are now available. It appears that pupils of the intermediate and grammar grades are the ones most favored in this respect. There are at least a half dozen group examinations for these grades which give evidence of meeting in a fair manner essential statistical requirements. There are also a number of examinations for the primary grades and at least two for college students. For practically the whole range of formal education we have therefore dependable means of intelligence measurement with more than a dozen separate group examinations available. Not all of these examinations are the product of the past year. Some made their appearance in 1919; and the initial work on most of them dates beyond a year ago. But 1920 will be noted as the publication year of a large number. Of particular importance is the completion and initial publication of the National Intelligence Tests in two batteries of five tests each with ten forms for each test.

Again, it may be said that the quality of intelligence tests has enormously improved. Five years ago at a meeting of the American Psychological Association in Chicago a speaker argued that no intelligence test yet devised had shown a reasonable correlation with school achievement and it was further pointed out that the several proposed tests showed but small intercorrelations among themselves. There are numerous tests now offered for use to which these strictures do not apply. Intercorrelations ranging from 0.60 to 0.80 are now fairly common and some reach to 0.90, and more. Correlations with good measures of school achievement are sufficiently high to furnish a dependable basis for prognosis. Continued and extensive experimental study has made possible the construction of tests meeting the essential criteria of discriminative capacity, reliability, significance, and adequate standards of comparison.

No claim is here made that our present tests are finally satisfactory. Sufficient evidence may be adduced to show that they are not, and that in behalf of scientific accuracy we must go on improving these tests, reducing the sources of error and determining the specific uses to which particular tests are best adapted. The direction which experimental work should take is fairly clear; certain of the important methods are known and important improvements in tests may be predicted.

Doubtless one of the most encouraging facts about our present use of intelligence tests is the increasing tendency to utilize test results in modifying teaching methods, and in school organization and supervision. In many American cities pupils are being classified on the basis of scores made in these tests, special classes are being organized for subnormal and gifted children, courses of study are being revised, methods of instruction are being changed, and in numerous other ways the assault is being carried against the deadening lockstep and uniformity into which current public educational method has fallen. That all of these changes are wise or that all will lead to a better education is not assured. But the pragmatic method implicit in such remedial work will furnish a corrective to unwise exploitation of the new devices. The path of progress lies through just such remedial efforts, even though the immediate outcome may not be entirely successful.

The current year is further marked by the effort to state achievement in relation to intelligence. The recent advance in the measurement of intelligence has called in question the use of achievement tests and other measures of school progress. Obviously, the formal school program should not be held responsible for low achievement due to the inferior intelligence of pupils; neither should it be credited with high achievement due to superior intelligence. The wide ranges of intelligence found in classes and schools of the same designation show that the direct interpretation of achievement tests does just this thing. Such test results are only interpretable when considered in the light of intelligence test results from the same pupils. Franzen and Buckingham and Monroe have proposed methods for calculating an "achievement quotient" which will adequately credit the school for capitalizing and paying dividends on whatever intelligence investment it may have. This definite recognition of native intelligence as a determinative factor in achievement is significant not alone for the interpretation of achievement test results, but also because it calls in question all those experimental studies in the learning of school subjects where no cognizance has been taken of the intelligence factor. Such studies—and their number is great—must all be done over before any satisfactory conclusions can be derived as to the efficacy of the other factors involved.

The improvement and extensive use of intelligence tests should not, however, obscure their limitations. To any careful experimenter it is obvious that tests of the type currently listed as measures of intelligence do not give us all the information about children that we need. If we admit that the so-called intelligence tests are really measures of intelligence then it is apparent that success in school work and success in life are not determined by intelligence alone. Other factors play a part and in many cases a determinative part in success and failure.

Several years ago two students working under my direction studied the characteristics of fifty men who were admittedly successful men and of fifty other men who were obviously failures in life. Each of the hundred persons was rated on a scale of eleven points for each of eighteen qualities. When the data were combined for the fifty successful men the results showed clearly that in the combined opinions of all the judges the quality most apparently conducive to success was *industry* which in the scale was defined as "thorough, persistent, painstaking, enduring" and the opposite of "lazy, sluggish, indifferent, superficial." The nine traits ranking next in order were "efficiency, attentiveness, loyalty, prudence, honesty, adaptability, sympathy, tactfulness, and cheerfulness." Only four of these nine traits could by definition be construed as in any sense equivalent to the current concept of intelligence. The results showed that the judges regarded such non-intelligence traits as industry, loyalty, honesty, tactfulness, sympathy, and cheerfulness as weighing heavily in favor of success and such other non-intelligence traits as self-assertion, pride, conceit, jealousy, quarrelsomeness, suggestibility, and intolerance as weighting individual capacity in the direction of failure.

Similar data are accumulating regarding the capacities of pupils to do the work of the school. A few individual cases may clarify the point. For the past five years, the pupils in the high school of the University of Minnesota have been examined with intelligence tests and the subsequent work of the pupils has been carefully followed and tested. In general the pupils scoring highest in the tests make the highest school marks but certain notable exceptions to this general rule stand out. Naturally they are of two sorts: those in which pupils with high intelligence scores

make low marks and those in which pupils with low intelligence scores make high marks.

When John Smith entered the high school in the autumn of 1917 he stood third in a class of 60 pupils in a series of intelligence tests which included the following: opposites, analogies, hard directions, verb-object, Trabue completion, and Thorndike reading scale Alpha 2. He was later examined with the Army Examination A in which he scored 325 points placing him easily in the upper 10 percent of high-school freshmen and the equal of many college students. On the Otis scale he scored 179 points. During the four years in high school, however, only twice has this boy achieved a mark as high as C+ on a five-point scale of marks. Of 27 marks so far received twelve have been D (the lowest passing grade), one has been C−, and the others have been C. Not a teacher who has had this boy in class but believes him capable of much better work than he does; but not a single teacher has been able to induce him to do it.

In marked contrast to this boy is the case of Mary Jones, a member of the same class, whose army test score on Examination A was 231, who scored average on the initial tests (ranking 23 in the group of 60 entering pupils), and whose I. Q. (Terman) was 108. Only five times in her four years of high-school work has this girl scored so low as B in an academic subject. Twenty-eight times in thirty-three her marks have been A and she is generally recognized by her instructors as the best student in the class.

Cases like these could be cited at length. Edgar Brown of the present freshman class whose combined scores on the Miller high-school test, the Terman group intelligence examination, and the Haggerty intelligence examination Delta 2 placed him in the lowest fourth of a class of 55 pupils, achieved rank 27 in the first quarter's school marks. At the same time Ernest Stuart, who stood in the upper tenth percentile in the combined intelligence tests, nevertheless fell to thirty-sixth place in school marks.

One might infer from cases like these that the tests are inadequate to measure intelligence. It may be granted that the inadequacy of the tests is a source of error and that a better test would give a more accurate prognosis of success or failure. The probabilities, however, lie in another direction, namely that intelligence in itself is inadequate to produce success. It is not at all probable

that a perfect measure of intelligence would give a perfect correlation with school success or with success in later life. A more accurate measure of intelligence would only render the inadequacy of intelligence more apparent for the simple reason that success is not quantitatively coterminous with intelligence but with intelligence in combination with other significant human traits not subject to evaluation by tests of the type currently used as measures of intelligence.

Before proceeding to consider these non-intelligence traits let us enforce this idea of the inadequacy of intelligence by further illustration.

John Ralph reported at a large city elementary school in October, 1919. At the time he was on parole from the state reformatory. His history disclosed a long record of truancy, incorrigibility, and lawlessness. He had at various times attended eight different grade schools, two opportunity rooms for subnormal children, a detention home for delinquent boys and the state reformatory. He was admitted to the elementary school at this time on a special permit because he had been refused admittance at his home school on account of his record of incorrigibility.

At the time of entrance he was 15 years of age and was classified in the VIIIA grade. He was well developed physically, but his scholarship was of exceedingly low quality. He was placed in the department classes without warning to the teachers. At the end of the week the teachers appeared in a body in the principal's office, pronounced the boy to be subnormal, stated that he was unable to do any of the work assigned, and asked the principal to have him sent to a special school.

The principal investigated his case and discovered that while "he could not write a good paragraph, nor solve a problem correctly, he seemed possessed of a shrewd practical mind. When he was given a problem he seemed to be wholly unfamiliar with the regular method of procedure, but would frequently give the correct answer off-hand and then in his written work proceed to prove his result. His English from the standpoint of mechanics would have shamed an average fourth-grade child; but although his spelling, punctuation, and sentence structure were very poor, his oral work showed that he could think clearly and logically."

The boy was passed into a group of pupils to be examined by a psychologist who was unaware of his history or school achieve-

ment. On the Terman test the boy showed an I. Q. of 122 and was clearly above average in mental capacity.

Special courses were arranged to give him an opportunity to get the fundamentals of English and mathematics which he had never acquired. He stopped his cigarette smoking in response to the principal's appeal and the next day whipped the biggest boy in school who chided him for "turning good." He repaired and washed up the principal's automobile and was allowed to drive the car on errands for the school. He was given many other executive responsibilities and proved himself competent in many ways. He graduated at the end of the year and while he was not so well grounded as other members of his class, he had attained a measure of self-respect, confidence in his teachers, and a good working attitude toward the school. He finished the course with a record for good behavior.

That cases of this sort are more numerous than might at first be supposed is evidenced by the results of the intelligence examination of 94 pupils in a so-called "opportunity school" in a large city. These pupils had been assigned to the opportunity school because for one reason or another they did not fit the schools to which they by geographical location belonged. They ranged in chronological age from 12 to 17 years. They scored in the Otis Group Intelligence Scale from 21 to 130 points. Three of these pupils whose chronological age was 13 years scored 126, 128, and 130 points respectively, which places them in Otis's genius class; fifteen others were rated very superior, or superior, and 25 were rated as normal. Yet all of these 45 pupils were succeeding so poorly in their school work that they had been segregated into this special school. There is certainly doubt that defective intelligence was the cause of irregularity in any one of the 45 cases.

One of the most complete demonstrations of the inadequacy of intelligence is to be found in the Report of the Laboratory of Social Hygiene at the State Reformatory for Women at Bedford Hills, New York. The report is the result of six years of investigation into the character and history of women delinquents in New York State and its conclusions are based upon more than 500 case histories from six of the state institutions which receive delinquent women. The data which were collected and studied

under the supervision of Dr. Mabel Fernald, Director of the Laboratory of Social Hygiene, are interpreted to show that "the average woman prisoner in New York State falls somewhat below the average individual in society in mentality and in economic efficiency." The authors point out, however, that this lower average does not imply that delinquent women are all from the lower end of the scale of intelligence. "The range of the delinquent group was found to be practically coextensive with that of the army group, our most representative sampling of the general population" and the authors further hold that their "data fail absolutely to justify the view expressed recently by certain propagandists that delinquency and defective intelligence are practically synonymous and that solving the problem of mental deficiency will solve the problem of delinquency."

Current psychological literature multiplies this type of evidence as to the inadequacy of intelligence. Madsen in a careful study in the South Omaha High School showed that the boys made in each class higher intelligence scores but lower school marks than did the girls. Colvin finds not only moderate correlations between school marks and Otis intelligence scores but that 60 percent of all the failures in the freshman class of the Reading High School were made by pupils who scored normal or above. Branson reports that certain pupils of apparently superior ability (Otis) fail to take hold of the work as well as the rest of the section. Buckingham finds intelligence tests insufficient for purposes of classification. Dickson finds one-fourth of his accelerated eighth-grade group making marks below the average for the school. Pressey notes that "if we wish to foretell success in school, we must obtain a measure of school attitude," and Terman insists "that mental tests should be supplemented by ratings on character traits and by educational tests."

In noting the progressive use of intelligence examinations during the past year it should be pointed out, therefore, that there has also been a growing recognition of the limitations of intelligence examinations and even of intelligence itself as a basis for prognosis. At the same time there have been efforts to supplement the measurement of intelligence with estimates and measures of the non-intelligence factors contributing to success. The direction of such supplementation is twofold. There is first the effort to find tests for special aptitudes as a basis for predicting

success in particular occupational fields. Thurstone's studies in telegraphy, Kitson's studies of proof readers, Stenquist's tests for mechanical information are illustrations.

In how far such special aptitudes exist apart from the factors which make for general intelligence is an unsolved problem. Sufficient evidence is available to warrant experimental work even though such experimentation promises to be more intricate and difficult than was the measurement of intelligence.

Along with emphasis on the non-intellectual factors as significant for prognosis and apart from the experimental work on special aptitudes there has been some attempt during the year to analyze such factors and to indicate the means for their evaluation. The most tried method for such evaluation is the so-called "rating scale" requiring personal judgments of non-measureable human qualities. For the initiation of this method as an instrument of psychological and educational research we are doubtless indebted to Cattell who used it not only in laboratory experiments but in his important work in evaluating the individual worth of American men of science. During the year Dr. Cattell has made an important addition to his method by developing "the probable error of a vote" taken in connection with the new forthcoming edition of this valuable work. The most ambitious effort to apply this method of estimating human qualities was the Scott rating scale for army officers used by the War Department during the recent war. The application of the method to educational work is not new but the year has seen numerous efforts to apply the rating method to the evaluation of capacities of pupils and to the qualities of teachers. The Detroit scale for rating teachers and the Pressey card for the rating of pupils make use of the Scott method.

By this method the judge makes up a separate rating scale for each quality to be considered. For the highest unit on the scale the judge chooses from among his acquaintances the person who in his opinion possessed that quality in the maximum amount. Similarly for the lowest unit on the scale a person with least amount of the quality is chosen. The midpoint on the scale is located by selecting a person who in the opinion of the judge stands midway between the highest and lowest units already chosen. Other points may be interpolated in any number between the ones thus fixed. By this means the judge constructs a rating

scale whose units are human individuals who are the personal acquaintances of the judge himself. Seemingly this device renders more concrete the estimation of human qualities and Scott has shown that in the rating of army officers the probable error in judgment is reduced by the use of this type of scale. Pressey reports "a coefficient of reliability (of agreement?) of 40" for different teachers' estimates of quality of "school attitude," but higher coefficients in the case of other traits.

Smith has published a scale for the rating of pupils which in his opinion has considerable validity and usefulness.

Kent criticizes previous teacher-rating schemes because they place so little emphasis upon the results which teachers achieve in the schools and recommends a rating scale in which the achievement of pupils may count for as much as 75 points in a total of 155 points which a perfect "professional" worker might achieve. Connor, who like Kent, proposes a teacher-rating scheme evolved out of the necessity of a school superintendent's supervisory and administrative duties, writes that "the means at hand for making scientific observations of the progress of pupils are scanty" but he believes such pupil-progress to be an important element in the rating of teachers.

In the Virginia survey I myself used a rating sheet upon which the teachers were asked to estimate the scholarship, intelligence, and industry of their pupils, using a five-point scale for each quality. A study of the results indicate generalizations somewhat as follows:

(a) Teachers using the same directions differ greatly in their ability to rate the intelligence of their pupils. (b) In general they make little distinction between intelligence and scholarship. (c) The judgment of the intelligence of a child is influenced by the school group in which he is found. Because of this an older pupil is usually rated above his intelligence and a younger pupil is rated below his capacity. In other words, teachers rate poor pupils better than they are and superior pupils as inferior to their real capacities. (d) A group intelligence test is a more generally dependable instrument for evaluating the intelligence of a pupil than is the teacher's judgment.

A consideration of all these rating devices leaves one with the sense of a problem discovered but unsolved. The variety and multiplicity of names used to designate traits to be rated

is bewildering. Some of these names are apparently intended to indicate psychological traits, others may be classed as social, others educational, and still others relate to the effect which an individual is able to produce on things or upon other persons.

Thorndike and Knight have shown by a study based on the use of the Scott army scale and the Boyce scale for rating teachers that there is a "constant error" in the use of rating scales due to the influence of "a marked tendency to think of the person in general," a tendency which suffuses ratings of special traits with a kind of halo. "The magnitude of this constant error of the halo seems surprisingly large," so large that the authors question the value of the ordinary type of rating scale for special qualities.

It would seem necessary to proceed by the method of experimental analysis to determine the group of traits important for rating and to determine the logical basis upon which correct rating scales can be based. It is further necessary to evolve by statistical methods the types of units and the number of such units for a usable scale, and finally the directions for the use of such scales must be definitely standardized.

While unhappy chaos marks the present use of rating scales, a glimmer of hope comes from another source. During the year Dr. June Downey has published additional results from her "scale for the measurement of the volitional pattern." This scale comprises objective tests, 12 in all, designed to measure such personal qualities as assurance, flexibility, speed of movement, motor impulsion, resistance, tenacity, coordination of impulses, freedom from inertia, motor inhibitions, care for detail, speed of decision, etc. The tests were first designed for individual examination but are being adapted to group use. The objective scores are platted in graphic form and the resulting graph is called "a will-profile." The author proposes these tests as a supplement to intelligence examination. She claims that "the will-profile has considerable general characterological significance and that it can be used to advantage not only in getting the general temperamental pattern of an individual but also in determining the specific combination of traits," and that in conjunction with intelligence tests it "affords in many situations a basis for conservative prophecy." In the author's three papers during the year she sets forth the results of the tests with varying groups and details a number of case studies. The

method has been used in the psychological laboratories of Cornell, Chicago, Carnegie Institute of Technology, and the University of Minnesota. It may develop important results in the measurement of character; but its chief claim to merit so far is that it initiates objective method in a field where our chief reliance hitherto has been upon subjective opinion.

These, then, appear to be the accomplishments of the year in the field of mental measurement: the extension of intelligence examinations in the public schools, the further development of rating scales for teachers and pupils, and the beginnings of objective measurement in the realm of non-intellectual traits. Certainly no other year, except that momentous first year of the great war, has witnessed more activity in this field. We are probably right in believing it to be a year of progress.

THE EFFECT OF CLEAR OBJECTIVES ON THE TEACHING OF READING

WILLIAM E. HAWLEY
Principal, Francis Parker School, Rochester, New York

In the primary grades reading is a vital subject. Teachers plan their work with exceeding care; minute matters are considered and given proper place; system is evidenced in every project; advice and information are keenly sought; every objective is clearly in mind and every road carefully charted. Pupils on the other hand feel the urge to success; they employ strict attention and endeavor; failure to succeed is serious, and success brings real satisfaction. In the upper grades, however, carelessness reigns; objectives are indistinct; paths of progress are unexplored; study of method is neglected. Imagination, concentration, judgment, and the memory are not stretched, although no other subject offers more opportunity for their development.

Thus the question arises, what motives can be supplied that will bring to the upper grades the vigorous mental efforts of the lower grades. First, the abilities in reading that are needed for further school work and for out-of-school activities must be more clearly determined. Second, definite measures of accomplishment, failure to meet which will bring shame, must be established. Third, more careful attention to method, materials, and to individual needs must be given.

The first of these needs is fairly met by the "Attainments in Reading" chart,[1] composed of material submitted by the teachers and compiled for the Rochester schools by Miss E. H. MacLachlan and a committee of teachers directed by Miss Mabel E. Simpson and Mr. J. P. O'Hern. Here we have, if they can be generally accepted, concrete, definite suggestions concerning what we should be attempting in the elementary grades. Thorough knowledge of this chart would indicate general appreciation of the difficulty involved in the satisfactory teaching of reading. In the light of it upper-grade reading cannot continue to be considered the easiest subject to teach. It challenges effort, preparation, persistence; and in these it is a brief as well as a decidedly welcome guide. Its use should bring emphasis and variety to the work and

[1] O'Hern, Joseph P. "The development of a chart for attainments in reading," *Journal of Educational Research*, 3:180–94, March, 1921.

should secure the attention of the pupils to achievement. An interest in the growth of ability to do the things demanded by the chart may thus be added to the mere interest in stories read.

Concerning the need for more definite standards; it is now possible for us to measure with a fair degree of accuracy the ability to read orally, the rate of silent and of oral reading, the ability to read short paragraphs and to react immediately under instruction and the ability to reproduce and to answer questions concerning long paragraphs. These abilities and others not here mentioned may likewise be determined by the use of informal tests. But, merely the ability to measure achievement in reading will not stimulate; the single giving of a standard test will not necessarily promote progress. Constant checking up of results for the purpose of charting progress so that pupils and teachers are aware of their needs is essential. By the use of constant checking the optimum achievement in reading may be established and this optimum will be demanded and obtained.

Concerning the need of more careful attention to method in the upper grades the general trend of effort may be accepted.

The meaning of the foregoing suggestions is explained and illustrated by the following description of an experiment attempted at the Francis Parker School during the fall of 1919.

In the spring of 1919 the Monroe, Courtis, and Gray reading tests were given to all our pupils. These tests showed that our pupils were satisfactory oral readers, that the rate of reading was about normal, but that there was a decided limitation in the pupil's ability to reproduce definitely and to answer questions correctly after reading. The fifth and sixth grades were shown to be relatively the poorest readers in these respects. So we decided to see what could be done to remedy the situation. We had three sixth grades in which were 109 pupils. We decided to concentrate our efforts for the entire term on securing progress in the quality of silent reading. We were influenced to this action not alone because our weakness was just here but also from a conviction that ability to reproduce underlies all success of the pupils in their future school work. Moreover, we believed that the ability to reproduce an article with something of its spirit, its logic and its idioms, cannot but be a valuable social and professional attainment. Argument often depends upon power to quote accurately from the thoughts of others. Reproduction requires

strict concentration and tenacious memory—two attributes necessary in all mental effort. So in thus restricting our teaching we were not unconscious of the other needs of reading but were merely expressing our confidence that the quality of silent reading was important.

Our sixth grades were departmentalized in arithmetic, geography, and English. We substituted reading for English and the pupils thus had forty minutes for reading each day.

Our first problem was to determine why pupils had failed in reproduction. Aside from any lack of strength which may have been present in the teaching we concluded that one factor was our general policy of discouraging the use of the words of the text and of constantly urging of pupils to "tell it in your own words." We cannot enter into much discussion of this point. We decided, however, that whenever an effective word used in the text could be remembered it should be used and that attempts should be made to remember some of the more expressive ones. In addition we tried to discourage the almost universal desire of pupils to bring into reproduction statements that did not appear in the text. Grade charts showing each pupil's records were shown and explained to the pupils and we did what we could to encourage them to improve. Daily reading for accurate reproduction, required for a variety of purposes, was practiced. Each week an informal

TABLE I. TWENTY WEEKS' PROGRESS OF 109 SIXTH-GRADE PUPILS AS SHOWN BY THE USE OF THE GRAY READING TESTS

	Oral Reading	Silent Reading			
		Rate (Seconds per 100 words)	Quality		
			Questions	Reproduction	Average
First test	48	42	40	20	30
Second test	53	40.9	59	27	43
Percents of gain	10	2.5	47.5	35	43
Fifth-grade standards	48	38.9			32
Sixth-grade standards	49	35.8			39

test was given which demanded the same activities as the Gray tests. These tests were printed, and copies handed to each pupil.

At the end of twenty weeks the pupils were retested by the Gray tests. About half of the pupils were fifth-grade pupils at the time of the first test. All were VIB and VIA pupils at the time of the second test. Scores were computed for those pupils only who were members of the school at the time both tests were given. (Only ten of the 109 pupils received a lower score in quality on the second test than on the first. Three made the same score on each test.) A summary of results is given in Table I. These results indicate that while the pupils were increasing their quality of reading 43 percent under intensive training for that purpose, the sole purpose of improving quality, they incidentally improved in oral reading five times as much as they normally would and at the same time made some improvement in rate.

The problem of attention to individual needs which we have mentioned as necessary for satisfactory teaching was met as follows: Twenty-one of the poorest readers among the 109 pupils in these sixth grades were selected from their records on the first Gray tests. They were then divided into four groups and arrangements were made to give each group about thirty minutes' special help in reading three or four times a week.

Our first concern with these pupils was to find out why they were unable to read. We concluded that poor reading might be due to any of the following causes: (1) physical defects; (2) lack of knowledge of English; (3) poor teaching and neglect; (4) lack of general intelligence; (5) special disabilities in reading. Examinations of the pupils showed certain physical defects in some cases but in only one case could these account for the difficulties in reading. Only one pupil was handicapped because of lack of knowledge of English.

We attempted to diagnose each case to find wherein each pupil was troubled. Let us digress here to remark that diagnosing the defects in reading is simple compared with the task of determining the cause of those defects. A doctor may take the pulse, measure the temperature and blood pressure, and pronounce his patient ill with a certain disease; but it may take long study to find out the cause of the disease. The symptoms of poliomyelitis and of cancer are easy to discover but the cause still defies detection. It is important, however, to observe that though the cause of the disease may not be known, relief or even cure may

often be effected. So in reading, inability to point out the cause of failure need not prevent us from applying measures of relief.

We had each of the twenty-one pupils of whom we have been speaking tested by the Stanford Revision of the Binet-Simon Test. We also gave them the Otis Group Intelligence Scale. These tests were applied in order to see to what extent low mentality might be a cause of trouble. Our reading tests and our judgment led us to suspect that this might be a cause in some cases. Yet low mentality was by no means a sufficient general cause. Table II shows that ten of the twenty-one pupils (all of whom were poor readers) had intelligence quotients greater than 100. Only six of them fell below 90, the lower limit of normality. Two of these were pupil No. 4 and pupil No. 17 (Table II). Yet these children made substantial progress in the interval between the tests. If their rather low I. Q.'s were taken as a full explanation of their poor scores in the first test, it would be difficult to explain how they could make such progress after the test.

To illustrate the attempts to study individuals we give the following description of three of the cases entered in Table II:

Case 13.—A girl whose age was 14-9; has a mental age of 11-1 and is therefore over three and one-half years retarded; born in Holland; lived in America four years; never learned to speak well; does not understand instructions easily; is not alert; has poor eyes and wears glasses. The doctor suggested that she be relieved from the regular gymnastic exercises because of physical difficulties. The girl has a sister in the school whose foreign birth could not be guessed and who does everything well. Teachers feel that her health and lack of mentality together with her lack of English are responsible for her poor reading.

Case 16.—A boy whose age was 11-1; has a mental age of 12 years; has no physical defects; has a fine home and unusually intelligent parents; does arithmetic well and succeeds in other subjects; no apparent cause for poor reading; reported that he knew he could never read and that he did not like to try; his father had tried to help him but without success; nevertheless he did well on the test. Asked at the end of the term how he could make so much progress, he reported that it was because of the general interest of the grade in reading and because his teacher helped him specially. He said that he had read his first book at home, the *Swiss Family Robinson*, that he liked it and that he was then reading a boy scout book.

TABLE II. THE RECORD OF TWENTY-ONE POOR READERS BEFORE AND AFTER TWENTY WEEKS OF INDIVIDUAL ATTENTION

Pupil No.	Sex	Grade[a]	Age at Beginning of Period (Years Last Birthday)	Mental Age at Beginning of Period (Binet)	Intelligence Quotient	Otis Scale Score	Rate[b] (Seconds per 100 words)		Quality[c] (Average of Reproduction and Questions)		Score in Oral Reading[d]	
							First Test	Second Test	First Test	Second Test	First Test	Second Test
1	G	VA	11	10.8	91	41	56	86	0	14	23	31
2	G	VA	11	11.2	98	64	27	32	6	29	56	55
3	B	VA	10	12.2	112	55	47	40	7	16	40	50
4	G	VIA	12	10.9	88	65	39	48	7	51	49	56
5	B	VIB	13	12.4	88	49	184	196	9	44	4	24
6	B	VA	10	13.8	124	60	43	37	9	47	54	56
7	G	VIB	11	15.4	128	71	26	35	9	35	40	44
8	G	VA	11	11.3	96	43	39	37	12	27	45	51
9	G	VIB	10	11.8	105	52	120	49	14	25	40	53
10	G	VIB	13	10.4	75	57	78	50	14	18	35	44
11	G	VIB	12	10.8	84	43	65	47	14	19	21	34
12	B	VIB	11	13.3	107	74	51	43	14	48	56	66
13	G	VIA	14	11.1	75	57	48	55	14	13	34	49
14	B	VA	12	13.7	111	81	28	27	14	21	50	51
15	G	VIB	12	12.8	98	74	59	36	16	27	36	40
16	B	VA	11	12	108	47	74	67	19	45	35	44
17	B	VA	12	10.8	83	71	40	37	19	36	35	43
18	G	VIB	11	12.8	110	69	60	53	20	36	49	55
19	B	VA	10	11.5	103	64	28	29	22	48	51	53
20	B	VIB	11	14.8	124	91	40	37	23	38	36	44
21	B	VA	13	12.7	93	58	114	76	41	51	34	40
Averages							60	53	14	33	39	47

[a] All pupils were promoted one half-yearly grade during the twenty-week period except No. 11 and No. 13. These two pupils were left back.
[b] Fifth-grade standard, 38; sixth-grade standard, 35. [c] Fifth-grade standard, 32; sixth-grade standard, 39. [d] Fifth-grade standard, 48; sixth-grade standard, 49.

Case 5.—A boy whose age was 14-1; has a mental age of 12-4 years; read with difficulty the simplest words; has very nearly been abandoned in so far as hope of teaching him to read is concerned; was interested in the records of the grade; very much wished to improve; for three years he had had difficulty in seeing, which had not been discovered; could do arithmetic if some one would read the problems but could not solve a single problem if compelled to read it for himself; his low mentality had evidently given him a poor start and his difficulties had increased from grade to grade. It seems that he is also the one out of the 109 pupils who has some special disability in reading. What the difficulty is we do not know. It evidently has to do with the rate of comprehension. At the end of the term he said that he had read a complete book, that he liked it and that he was beginning to read rapidly enough to enjoy the story. He added, "My parents are surprised to hear me read from the newspaper."

The details of the devices used for improvement cannot be given here. The teachers employed word drills, sandtables, progress charts, home reading of assignments with written reproductions and weekly tests. The pupils made joke-books, asked questions, read for visitors, dramatized and discussed their own achievements. Emphasis and variety were the watchwords. The error in attempting to judge reading merely by hearing pupils read can be judged by this study.

In previous years our efforts had been expended largely upon oral reading without attaining a high degree of success. When the emphasis was shifted to silent reading for meanings, the oral work improved decidedly.

One conclusion to be drawn is that the excuse of pupils being "poor material" is a poor one. A second conclusion is that if pupils have definitely established standards toward which to work, pride will enter in and the work will improve. On the other hand, while mental and physical defects are to some extent the cause of poor reading ability, nevertheless, standards, frequent testing, and special help from the teacher will result in better reading in spite of these defects. Still another conclusion, and perhaps the most important, is that an effort to improve the quality of reading in a grade will result in a general improvement in oral-reading attainment and ability to reproduce.

AN EXPERIMENT WITH THE OTIS GROUP INTELLIGENCE SCALE IN THE NEEDHAM, MASSACHUSETTS, HIGH SCHOOL

ROSCOE L. WEST
Superintendent of Schools, Needham, Massachusetts

Group intelligence tests are being used for various purposes: to divide a school or class into divisions according to ability, to aid in guiding a pupil in his selection of courses, to give teachers information concerning pupils, and so on. One object in the experiment described here was to see if these tests could be used to stimulate pupils capable of doing good work to live up to their capabilities. It was felt that this would be a very practical benefit and that if the tests could be used for this purpose they would be of constant value in improving the standard of work of many pupils.

On December 29, 1919, the Otis Intelligence Scale (Form A) was given to 173 high-school pupils divided as follows: grade XII, 26 pupils; grade XI, 27 pupils; grade X, 46 pupils; grade IX, 74 pupils.

In order that the test might be given simultaneously, the whole school was divided into five sections. These sections were given the test by the superintendent of schools, the high-school principal, and three teachers. The test had first been given to the teachers by the superintendent; and the procedure and methods had been discussed so that conditions might be as nearly uniform as possible.

Medians and ranges in the Otis Group Intelligence scores (Form A) are indicated in Table I.

TABLE I. RESULTS OF OTIS GROUP INTELLIGENCE SCORES, FORM A

Grade	Median	Range
XII	164	209 to 90
XI	159	183 to 104
X	143	205 to 61
IX	129	191 to 70

In order to secure an accurate record and to make some comparisons, Form B was given on January 16 with the medians and ranges as indicated in Table II.

TABLE II. RESULTS OF OTIS GROUP INTELLIGENCE SCORES, FORM B

Grade	Median	Range
XII	182	206 to 109
XI	175	191 to 121
X	160	210 to 80
IX	151	200 to 82

The records for Form B are distinctly higher than those for Form A, due probably to the familiarity with procedure acquired in the first test. The average difference between the medians of Forms A and B is 18 points. The relative positions of the pupils were, however, about the same in the two tests as shown by the correlations between Form A and Form B in Table III.

TABLE III. CORRELATIONS BETWEEN FORM A AND FORM B

Grade	Correlation Coefficients
XII	0.92
XI	0.89
X	0.95
IX	0.91

These high correlations indicate that the two forms test the same things.

Another set of figures (shown in Table IV) was made up to compare the records of pupils of various ages in the different grades.

Table IV shows clearly that the youngest pupils in the grade made the best records. With one or two exceptions (where only one pupil is in a group) the scores decrease as the age increases. Upon examination of the scores of pupils of the same age in different grades it is found that the highest scores in each case are found in the highest grades. For instance, the sixteen-year-old pupils in grade XII score 168, while those who are 16 in grade XI score 152, in grade X, 139, and in grade IX, 115. These results indi-

TABLE IV. COMPARISON OF RECORDS AND AGES IN GRADES XII-IX, FORM A, OTIS SCALE

Age	Grade							
	XII		XI		X		IX	
	Number	Average Score	Number	Average Score	Number	Average Score	Number	Average Score
13	.....	.....	.....	.....	.....	.....	8	137
14	.....	.....	.....	.....	1	140	27	135
15	.....	.....	2	174	13	155	26	118
16	3	168	9	152	20	139	10	115
17	12	168	14	151	10	125	3	113
18	8	156	1	145	2	120	.....	.....
19	2	127	1	104	.....	.....	.....	.....
20	1	135	.....	.....	.....	.....	.....	.....

cate that the pupils of high intelligence have been able to complete the elementary school in less time than those of lesser intellectual capacity. The retardation of the older pupils is evidently accompanied by a certain lack of intelligence.

The next problem considered was the correlation between the intelligence tests and the teachers' marks. For this purpose the marks given during the fall term (two periods of eight weeks each) were considered. When the average of the Otis scores (Form A and Form B) and the marks were related, the correlation coefficients were found to be as follows: grade XII, 0.66; grade XI, 0.59; grade X, 0.55; grade IX, 0.58.

These correlations are fairly high and show a marked tendency for the distribution of teachers' marks to arrange themselves according to intelligence as shown by the tests.

We then prepared diagrams which showed how each pupil ranked in relation to the other members of his class both in the intelligence test and in his marks. Naturally some pupils were higher in marks than in intelligence and some lower. We felt that we did not need to worry about the pupils whose work graded higher than their intelligence. The most outstanding cases of this kind were discussed to see if the teachers were giving too high marks. In each case the superior rating in school work was interpreted as being due to hard, conscientious application on the part of the pupil.

TABLE V. RELATIVE RANKING OF TWENTY-TWO SELECTED PUPILS IN THE INTELLIGENCE TEST AND IN MARKS

Grade	Pupil Number	Otis Rank in Class	Mark Rank in Class	Reason for Difference
XII	1	2	13	Outside work
	2	5	17	Slowness
	3	17	24	Absence due to social activities
XI	4	3	14	Outside interests and lack of study
	5	5	20	Diffidence and lack of knowledge of how to study
X	6	4	18	Lack of study
	7	5	30	Absence due to home conditions
	8	17	37	Uninterested, perhaps due to poor health
	9	10	22	Too interested in social activities
	10	18	31	Absence due to social activities
	11	21	35	Absence due to home conditions
IX	12	1	15	Lack of ambition and application
	13	4	29	Doesn't know how to study—can't follow directions
	14	13	53	Chronic laziness
	15	7	26	Slowness in written work
	16	5	22	Laziness—frequent absence
	17	9	39	Indifference
	18	17	58	Absence due to home conditions
	19	18	63	Doesn't try—seems immature
	20	34	57	Wastes time in dreaming
	21	12	27	Lack of ambition
	22	25	42	Immaturity

We then selected the most glaring cases of pupils who in marks were below their relative ranking in the intelligence test. Twenty-two of the thirty-one pupils selected for study remained throughout the year so that comparisons with their marks during the last two periods of the year could be made. These twenty-two cases were discussed in a meeting of the teachers to find out why the teachers thought they were not doing work of which they were capable. Table V shows the relative ranking of the pupils, both in the intelligence test and in marks. It also gives the principal reasons assigned by the teachers for poor work.

The amount of differences in ranks in the classes and the average difference is given in Table VI.

TABLE VI. SUMMARY OF DIFFERENCES IN RANKS

Grade	Pupils	Total of Differences in Position
XII	3	30
XI	2	26
X	6	98
IX	11	286
Total.....	22	440

Average difference in position: 20

Table VI indicates, therefore, that these 22 pupils averaged 20 positions lower in their ranking by marks than by the intelligence test.

It is evident, however, that the opportunity for displacement in relative ranking in a small class is not equal to that in a large class. For example the margin separating the eighteenth from the nineteenth pupil in a class of 26 is larger than the margin between the eighteenth and nineteenth pupils in a class of 74. The figures should, therefore, be reduced to a common basis if statistics covering them all are to be of value. It would seem to be easiest to reduce all classes to a basis of grade IX in which there were 74 pupils. To do this, we may simply multiply the rankings given in Table V by the quotient found by dividing 74 by the number in each class. These quotients are 2.8 for grade XII, 2.7 for grade XI, and 1.6 for grade X. The rankings adjusted by this method give the differences in positions shown in Table VII. Ac-

cording to this adjustment the differences in position were such that if each of the three smaller classes had been enlarged to 74 pupils, without changing its characteristics, each pupil would have averaged 27 ranks lower in marks than in intelligence score.

TABLE VII. ADJUSTED DIFFERENCES IN POSITIONS

Grade	Pupils	Adjusted Differences in Position
XII	3	84
XI	2	70
X	6	157
IX	11	286
Total.....	22	597

Average difference in position: 27.1

After these facts were compiled a letter was sent to the parents of each of these 22 pupils explaining the intelligence test, the difference in standing between the pupil's ranking in this and in his marks, and the reason assigned by the teachers for this poor showing. The parents were asked to aid the school in bringing these pupils nearer to their proper standing as shown by their intelligence records. Table VIII shows the relative standing of the pupils in the Otis test and in their marks for the first two periods and for the last two periods of the year. It should be noted that the rankings are given as *adjusted to the class of 74 pupils.*

Sixteen pupils, as shown in Table VIII, gained in position from two to thirty points. The total gain was 176 points or 11 per pupil. Six pupils showed a loss of 53 points or 8.8 per pupil. For the entire twenty-two the net gain was 123 points or 5.5 per pupil.

This investigation shows that, of the pupils making the largest gains, numbers 7, 9, and 10 were pupils whose poor work seemed to be due to home or social activities, matters which could be corrected. Number 5 was reported as lacking knowledge of how to study. Special attention was given to aid her in her study habits with a resulting gain of 30 points in her standing. Number 15, who was reported as slow in written work, had probably been marked unfairly because some teachers did not appreciate his real ability. The four pupils who made the largest losses, numbers

TABLE VIII. RELATIVE STANDING OF PUPILS IN OTIS SCALE AND IN MARKS FOR FIRST AND SECOND PERIODS

Grade	Pupil	Otis Rank	Ranks on Teachers' Marks		Gain	Loss	Teachers' Marks in Percents		Gain	Loss
			First	Second			First	Second		
XII	1	6	36	34	2		83.5	85.2	1.7	
	2	14	48	31	17		78.8	86.1	7.3	
	3	48	67	70		3	73.3	77.7	4.4	
XI	4	8	38	35	3		79.3	77.1		2.2
	5	14	54	24	30		74.4	80.6	6.2	
X	6	6	29	27	2		76.5	75.2		1.3
	7	8	48	19	29		69.7	78.5	8.8	
	8	27	59	56	3		63.6	62.6		1.
	9	16	35	21	14		74.8	78.4	3.6	
	10	29	50	31	19		71.0	73.	2.	
	11	34	56	53	3		68.4	66.7		1.7
IX	12	1	15	30		15	82.5	77.3		5.2
	13	4	29	27	2		79.5	80.1	0.6	
	14	13	53	61		8	70.5	63.7		6.8
	15	7	26	10	16		80.2	87.1	6.9	
	16	5	22	18	4		82.7	83.1	0.4	
	17	9	39	46		7	75.5	72.6		2.9
	18	17	58	60		2	69.	64.7		4.3
	19	18	63	52	11		67.2	69.7	2.5	
	20	34	57	47	10		70.	72.2	2.2	
	21	12	27	45		18	80.	73.1		6.9
	22	25	42	31	11		74.2	76.7	2.5	

12, 14, 17, and 21, were reported as "lacking ambition," "lazy," and "indifferent." Evidently these drawbacks were not corrected by the methods used as the last standing was worse than the first. These special notes show how the tests can be used to follow up individual cases.

The average of all the pupils for the first period was 74.7 percent and for the last period, 75.5 percent; a gain of only 0.8 percent. This is, however, of little significance unless we know how the marks of all the other pupils compared. The figures given as to relative standing in the class are of more significance. These

TABLE IX. CHANGES IN TEACHERS' MARKS BETWEEN THE FIRST AND SECOND PERIODS

Grade	First Mark	Second Mark	Gain	Loss
XII	78.5	83.0	4.5	
XI	76.8	78.8	2.0	
X	70.7	72.4	1.7	
IX	75.6	74.6		1.0

figures show that the sixteen pupils who gained made a considerable increase. Of the six who lost, two brought the record down very much by losing 15 and 18 places respectively. It seems, therefore, very safe to say that intelligence tests will be found useful as one factor in stimulating pupils to the best work of which they are capable.

The results of this experiment may be summarized as follows:

1. There was a very high correlation between Forms A and B of the Otis Scale.

2. There was a fairly high correlation between the Otis test and the teachers' marks.

3. There was a very definite tendency for the youngest pupils to make the best records.

4. A considerable gain was produced by using these tests with 22 pupils to bring their work up to a standard nearer that shown by the results of the intelligence test.

RAYMOND: A CASE STUDY ILLUSTRATIVE OF THE DIAGNOSTIC LIMITATIONS OF A QUANTITATIVE SCORE

BERNICE LELAND
Marr School, Detroit Teachers College

The case of Raymond is presented here as suggestive of a class of children whose departure from the accepted standards of behavior and proficiency makes them conspicuous problems in the schoolroom. A surprising discrepancy is observed between the demonstrated ability of these children to learn school subjects by the usual school methods and an intelligence quotient, determined on such a scale as the Binet. The educational prescription can be most effectively determined by the analytical consideration of the child in his efforts to learn, such a study supplementing the tests and completing the diagnosis.

Raymond has been under careful observation and special instruction for twelve school months, the educational procedures varying as one failed and another seemed feasible. He is now (September, 1920) eight years and nine months old and has been in school without long absences since he was five. His intelligence quotient (Binet) in June, 1920, was 105.

So difficult and uncertain, however, is the process of learning school subjects that he has made comparatively little progress. In June, he scored *2* on Haggerty's Achievement Test, Sigma 1, and *1* on Test 2. When school closed he had read with great effort some fifty pages of a first reader. Now, three months later, he can scarcely read acceptably the primer which preceded the other book and which he must have first encountered over two years ago. He writes very well, forming the letters with care. He does not copy accurately, however, omitting and substituting letters. He has acquired a few number facts; but he is not keen at problems, answering impulsively and without much deliberation. His spelling vocabulary is very limited and, like all of his school performances, erratic-fair today, impossible tomorrow. He is fond of stories, but takes little part in language activities. As he talks, his lips tremble and his voice is forced as though his jaws were set. He is not timid.

An observer characterizes him at once as "queer looking." His ears are prominent and his thin little body supports a head

disproportionate in size and slightly brachycephalic. His blue eyes are very wide open and he works feverishly.

There are days when the dynamics of an idea carry him into appropriate action with such sudden discharge that he is up and gone before one can stop him. As he dashes into the hall, his teacher calls, "Raymond, it is nearly recess time. Shall we wait for a few m'nutes?" He always agrees readily, and settles down to his work without the slightest display of discomfort. Again, he seizes his pencil, making a precipitate dash for the sharpener In the midst of what appears to be close application to his assignment, the idea of drawing pictures, which he does with fair skill, or of making paper furniture perhaps, gains ascendancy. Raymond discards all else and is presently seen suiting the action to the idea.

A situation requiring organization and the focusing of certain competencies for the achievement of an unaccustomed end is a very difficult one for Raymond. He can neither persist along a certain line for any considerable time nor can he muster his forces efficiently at the start. Given a paper, upon which appear two or three simple questions and the number of the page upon which the answers may be found, he is occupied for some time; but he accomplishes nothing, although it is determined that the task is not beyond either his reading ability, meager as it is, or his comprehension. He is always "just going to do it."

This mental ineffectiveness has resisted all attempts at correction. It results in an exceedingly limited daily output, and has been an emphatic obstacle to his educational progress. There is no malice in his conduct. He does the best he can, and he is pitifully conscious at times of his failures.

The retentive phase of Raymond's memory is conspicuously disturbed. "Where is your book, Raymond? Boys in this class need their books every day, you know." After a search for the missing reader, he laughs and explains that he had left it "out in the hall on the top step." A futile trip to the stairway follows. "You must find your book, Raymond, now." Bewildered and excited he dashes out of the room and presently we see him pass the window on the run. The book was at home and when he finally remembered where it was he had intelligence enough to go after it without consulting anyone. This is a typical reaction. School material, notes or forms to be signed by his parents and

returned, rubbers or a sweater to be taken home, all slip through his memory like water through a sieve. He even forgets that he has not finished his assignment of work and merrily engages in some other pursuit after an interruption for a reading lesson.

This unfortunate lack of retentiveness displayed in the every-day affairs of his life is emphasized for symbols, particularly words. His associations are weak and any system intended to teach a reading vocabulary by rote learning, therefore, holds no advantage for Raymond. When he comes upon a word which he cannot recognize, he cannot remember where he saw it before, where he heard it or read it last, nor where in his book to look for the sentence if he could keep in mind such a prop to lean upon. We have tried writing the word, saying it aloud at the same time, spelling it, and calling attention to details such as "two t's in the middle" of the word. We have used pictures, alphabet cards, word cards, sentence strips, flash cards, and what not to stimulate in every way for a given word. We have worked on phonetic words, and, as he has no difficulty with blending, he gets along fairly well for a time. Then, without warning, he may begin the application of phonics to every word he sees—and English is not a consistently phonetic language. All our efforts to reinforce weak visual imagery avail us nothing. Raymond cannot read. Furthermore, he seems to have an "idiosyncracy" for certain words, particularly "spring." Special emphasis has again and again been given it, and he has had months of experience with books in which it occurs often. If the word were presented to him at this moment, the chances are about even that he would not know it. Doubtless faulty attention explains some of this, but some special disability adds to the sum total of his handicaps.

It is interesting to note the father's statement. He says that Raymond is "just like" himself. He is a broker and has resorted to a system of "memory training" because he could not remember his price lists. He undertook to teach his son to read by this same device. He admits that the plan was a failure.

Visual imagery not concerned with symbols seems to be adequate. Raymond possesses a certain manual dexterity which enables him to construct paper toys nicely without a copy, to build with blocks, and to handle tools for simple woodwork. But his resources are soon exhausted and he resorts to building the same thing day after day. Last spring he made pin-wheels every afternoon for days after we had each made one to take home. At

last we aimed his ideas in another direction and he became an ardent builder of garages.

On the Binet scale in June, 1920, he failed on the following tests: eight-year—counting backwards; nine-year—date, making change, repeating four digits backwards; ten-year—all except the absurdities. One-half of the design (a) was correct. (In September, 1919, the other half of the design was correct.) His plan for the ball and field was a spiral.

His basal age was seven years; his mental age was eight years and nine months. This quantitative expression of results is not an adequate statement of the situation. The emotional instability which Raymond exhibits is not a determining factor in the test, nor is the memory disability. His attention is persistent enough for each test, and his trainability for facts is high enough so that in his nearly nine years of existence he has acquired enough information to carry him through. It requires a qualitative analysis of behavior to set forth the facts in the case.

We do not know the cause of the boy's condition, but we have a theory. We know that he was not physically perfect at birth, a slight operation being necessary before micturition could be accomplished. We know that he was frail, "almost had the rickets," and did have acute adenitis when one year old.

After a Herculean effort on our part, his parents were persuaded in November, 1919, to have his tonsils, which had been badly diseased for a long time, removed. At the same time, the dentist extracted eight ulcerated teeth. Raymond was then seven years and ten months old. We were encouraged by the fact that he was calmer after this, and that *for the first time in his life*, he was persuaded to eat solid food. We very soon found, however, that he resisted learning just as before, and that he was equally unstable at times; though it is true that such times were somewhat less frequent.

Six months after this experience, Raymond weighed 48.5 pounds, was 48.8 inches tall, and a physician diagnosed his case as one of vegetative hypotonia. After four months of delay, the parents have not yet adopted the vitamine feeding prescribed, and we believe that the boy's nervous system is so devitalized by disease and lack of proper nourishment that he presents a clinical picture of irremediable mental defect. If there ever was a time when the nervous mechanism could have been aided to make a harmonious and consistent contribution toward standard behavior, we fear that time has passed.

RATING THE GENERAL APPEARANCE OF PRIMARY NUMBER WORK

ERNEST C. WITHAM
Superintendent of Schools, Southington, Connecticut

Professor Bagley says in his book on classroom management that: "One of the most accurate indices of a teacher's efficiency is the character of the papers and of the black-board work that his pupils produce."[1] The same author also says that: "Neat papers and neat blackboard work involve a number of little, specific habits, and each of these must be taken up and drilled upon apart from the others, and in some measure, apart from the content or thought of the work that is being expressed."

Some teachers are apt to be concerned entirely with the accuracy of number papers and to neglect their appearance, provided the work is correct. Superintendents, principals, and, for that matter, teachers themselves have only a limited amount of time to devote to this detail. However, it is an important one and it must receive enough attention to obtain good results. Many teachers require neat papers, but in any large system of schools there are always some who in this respect as well as in others need considerable supervision. To remedy this fault as far as possible and with least outlay of time and effort, the chart which is here described was devised for rating the general appearance of number work.

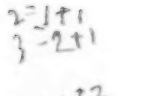
This particular chart is not offered as something scientifically accurate—it was not constructed in the shade of a great university, nor is it the result of work with thousands of pupils, hundreds of assistants, and some professors. Nevertheless, it has proved to be practical, because it gets results; and it is offered as a suggestion to other busy school officers. If some such device as this is occasionally in evidence, work of the type shown in specimens 1, 2, and 3 will be very rare.

For the benefit of those who wish to construct a general appearance chart, it is suggested that they gather up the number papers of from five hundred to one thousand pupils in grades I to V.

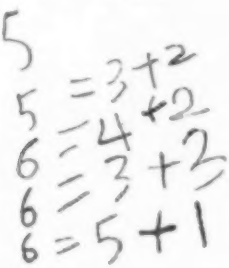
[1] Bagley, W. C. *Classroom management; its principles and technique.* New York: Macmillan Company, 1916.

CHART FOR RATING THE GENERAL APPEARANCE OF PRIMARY NUMBER WORK

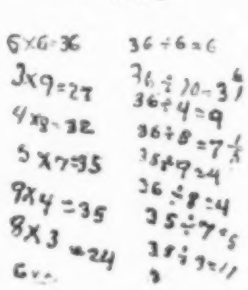
No. 1



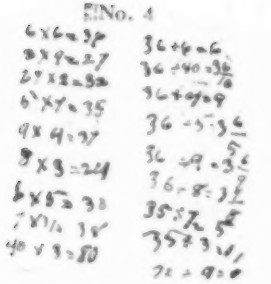
No. 2



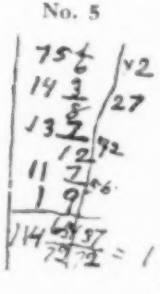
No. 3



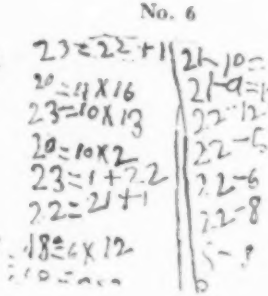
No. 4



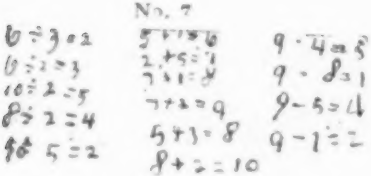
No. 5



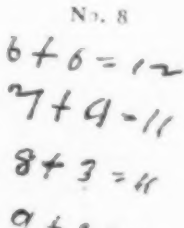
No. 6



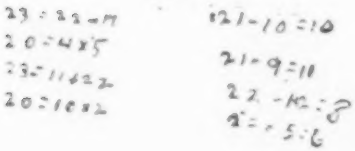
No. 7



No. 8

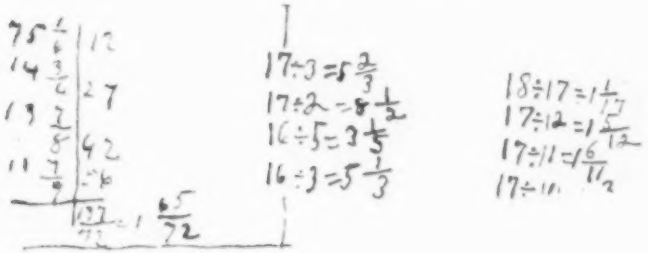


No. 9

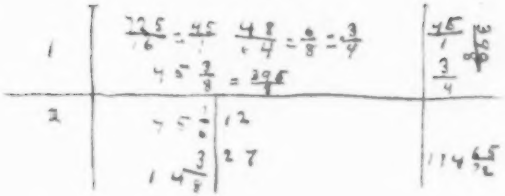


No. 10

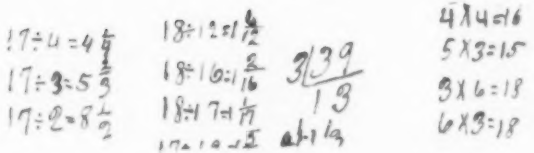
No. 11



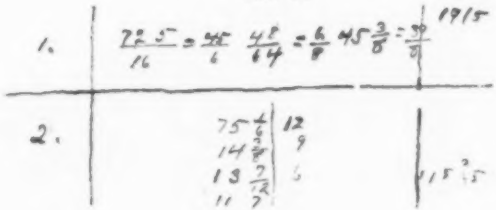
No. 12



No. 13



No. 14



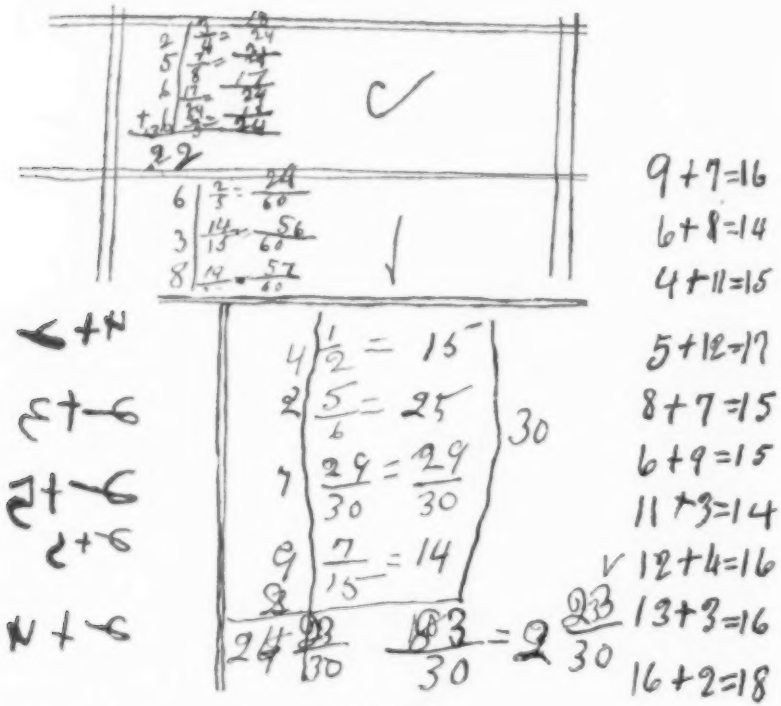


Figure 1

This should be done without previous warning to pupils or teachers. From these papers one will find an abundance of material with which to construct a chart. The papers should be carefully divided into fourteen or fifteen groups, according to excellence of appearance. One sample should then be selected from each group and cut down to a size convenient to use on a chart. The samples should then be fastened temporarily with paper clips to separate pieces of cardboard. Each specimen, when thus mounted, should be identified by a letter. In Figure 1, for example, the specimen is marked F. One may now rearrange the various specimens in what is judged to be the correct order, from the poorest to the best. It is desirable to have the judgment of several teachers in regard to arrangement, first, because many opinions are better than one and second, because it is wholesome for the teachers to have a hand in such matters.

A good way to secure the judgments of teachers is to hectograph copies of Figure 2, giving one copy to each teacher with instructions to record the arrangement of the samples which she thinks is best. She can do this by writing the identification letters in the blank spaces, thus producing a result like that shown in Figure 3. One may then figure out the correct order in which the specimens should be placed on the final chart. When this is done, one has something that is as useful as the "Graphometer" eye-test cards or any other chart commonly used in school work.

Figure 4 shows the ratings of 128 papers according to the "Chart for Rating the General Appearance of Primary Number Work." The pupils were distributed from the first to the fifth grade. This graph vividly shows the great range in quality of the papers, judged only from appearance. The distribution is nearly as great as are typical distributions for accuracy. For example, Figure 5 shows a distribution according to accuracy. The material is obtained from the use of Courtis Standard Tests in Arithmetic, Series B, Addition.

The irregular line on Figure 4 shows the median scores in general appearance for the various grades. Grades I, III, and IV are probably about normal. Collection was made in the second grade on a day when the teacher had tried to secure extra good papers. In the case of grade V, the work is decidedly uneven. There is an apparent lack of responsibility on the part of the teacher and pupils for the general appearance of the written work. No amount of explaining by the teacher that this was an unusual day is sufficient excuse for this slovenly work.

Figure 6 is a diagram of the actual relation of general appearance to accuracy in a set of fifth-grade papers. The figures at the bottom denote the eight different percentage ratings assigned to the papers by the teacher. The papers were marked on accuracy alone and the examples were either right or wrong. Effort did not count. It will be noticed that many of the best papers have a very low rating for general appearance.

The data here presented are not sufficient to warrant us in drawing definite conclusions. They do, however, point out the fact that the general appearance of primary number work is not necessarily closely correlated to accuracy. In many cases habits of neatness must be developed aside from those of accuracy. It

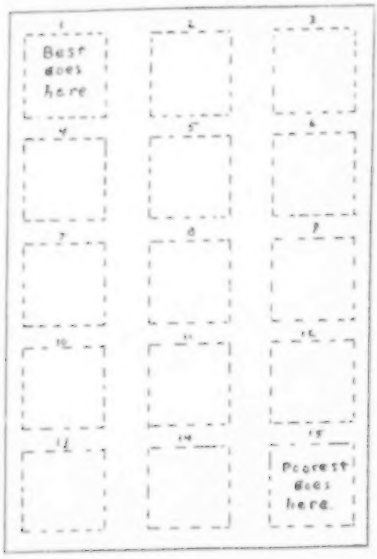


FIGURE 2. A SUGGESTED FORM FOR USE IN SECURING TEACHER'S JUDGMENTS

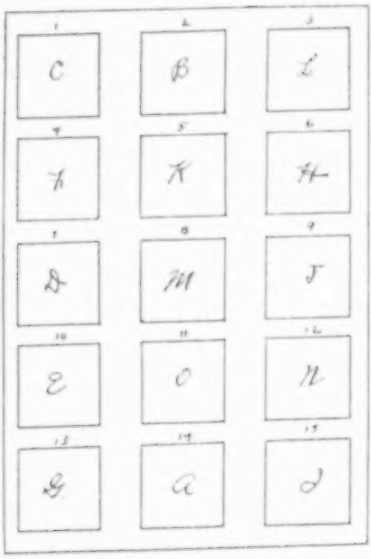


FIGURE 3. THE SUGGESTED CHART (FIGURE 2) AS MADE OUT BY A TEACHER

GRADE	GENERAL APPEARANCE													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	••	••• •••	••• •••	•• ••	•	•	•	••						
2		•	••	••	••	••	•• •	•• ••	•	••• •••	•• •	•• •	•	•	
3					•	•• •	•	•	•• •	•		•		
4						•	•	•• •	•	••• ••	•••	••• •••	•	•
5		•	•	••• ••	•• ••	•• •	•• ••		•• ••	••• •••	•• •	••	••	

FIGURE 4. DISTRIBUTION OF PAPERS ACCORDING TO GRADE AND GENERAL APPEARANCE

may be that the higher pupils go in the grades the greater the correlation between neatness and accuracy. In the eighth grade it may be very high, but no close relationship exists in the case of the pupils we have considered.

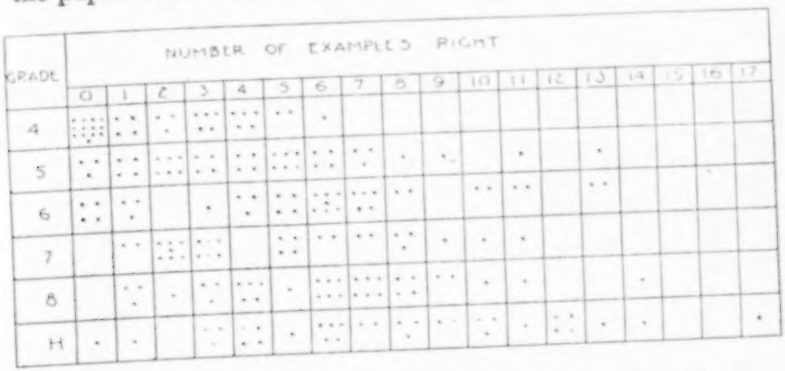


FIGURE 5. AN EXAMPLE OF DISTRIBUTION ACCORDING TO ACCURACY —COURTIS STANDARD RESEARCH TESTS IN ARITHMETIC SERIES B. ADDITION

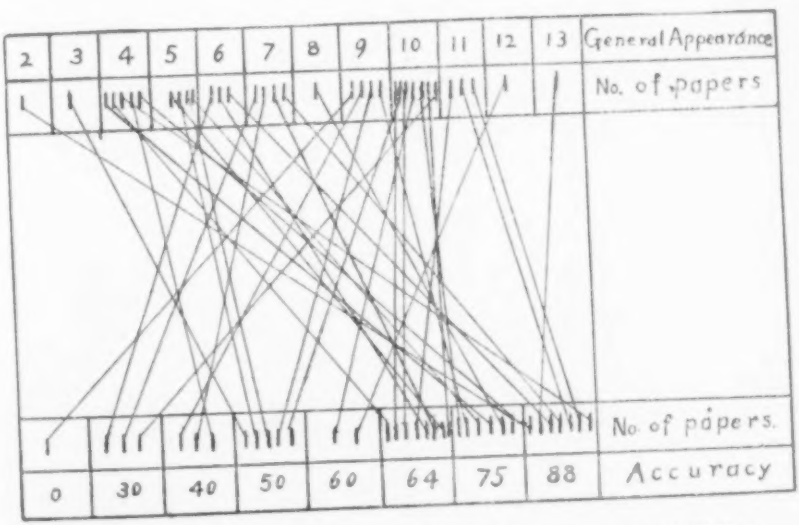


FIGURE 6. RELATION OF APPEARANCE TO ACCURACY IN FIFTH-GRADE PAPERS

METHODS OF INTERPRETING RESULTS IN THE CLEVELAND ARITHMETIC TESTS

JOSEPH PETERSON

George Peabody College for Teachers, Nashville, Tennessee

In a number of respects the Cleveland Survey Arithmetic Tests are the best we have for determining by group methods the relative strengths and weaknesses of a class, a school, or a school system in fundamental operations. These tests, however, are fifteen in number, and when one comes to the practical necessity of organizing the results for the several grades of the different schools of a system, one becomes painfully aware of the need of some weighting scheme for the condensation of the data. Only by such condensation can a report show to the average citizen, who votes school funds, and to the educational people who use them, just what each should know about the schools. Several survey reports of large systems tested with the Cleveland tests have appeared, and all seem clearly to reveal this need. Many tables and figures are found in these reports, but in no case does a final summarizing graph show just the status of a system in all the fundamental operations together.

An attempt was made to present such a showing after the Cleveland Survey Tests had been given to the children in the B grades of the Peabody demonstration school.[1] The median scores in each of the fifteen tests as obtained at Cleveland[2] and Grand Rapids[3] were accepted as norms, and the corresponding medians for the demonstration school were expressed as percents of these norms. Thus a percent of more than 100 indicated that for the grade and test in question the result at the demonstration school was better than the norm, while a percent of less than 100 indicated the opposite. Table I shows these percents, those less than 100 being in heavy type.

[1] In this school, which is connected with the George Peabody College for Teachers, educational surveys—and usually mental surveys also—are made annually.

[2] Judd, Charles H. *Measuring the work of the public schools.* Cleveland, Ohio: The Survey Committee of the Cleveland Foundation, 1916. 290 pp.

[3] *School survey, Grand Rapids, Michigan,* 1916.

Table I shows clearly certain weaknesses which would probably not have been pointed out for correction if certain tests of less diagnostic value had been employed. For instance, it is evident that in the demonstration school disproportionate emphasis has been placed on the simpler aspects of addition—adding a few single-digit numbers. At the same time the more complex processes in this operation have been neglected. Pupils have not been drilled enough in retaining the accumulating sums of long columns and in the process of carrying to other columns. Tested by the Courtis tests, the school would doubtless have shown a general weakness in addition which the teachers would have tried to correct by more emphasis on the very thing now overdone, according to our norms—namely, quick addition of short columns without carrying. Similar benefits of more careful analysis than is usual are also evident in some of the other operations, particularly in multiplication and in division. The more complex processes in these operations need more drill.

But this method of showing results is capable of improvement. Absolute scores are often of little significance unless norms are shown not only for the grade tested but also for the grades above and below it. Such norms are unobtainable in cases of tests—as the Monroe Standardized Reasoning Tests and the Monroe Standardized Tests in Silent Reading—that are different for upper and lower grades. The justification for the use of such tests is the practical difficulty of getting tests with enough progressively harder steps of equal size to discriminate sufficiently well between the abilities of any two consecutive grades, from the lowest to the highest—a difficulty clearly shown for example, in the Trabue Language Scale. Upper and lower norms, as well as norms of the grade tested, shown either in tables or graphs, or both, add considerably to the interpretative value of reports, both for individual and for group diagnosis; and it is not difficult to show such norms when all grades take the same test.

It is also desirable, if possible, to interpret the deviation from norms or standards in terms of years of progress. As our schools are organized a year's work is a natural unit of measure, and even the better trained educators—to say nothing of parents and the general citizen who are always to be kept in mind—think of retardation or acceleration in terms of years (or grades). This

unit, moreover, is the one most successfully used in the conventional tables and reports. Though progress from grade to grade is not always the same in absolute amount, it is the same in the practical sense that it always takes the same time to secure it.

Finally, it is obviously desirable for facilitating a general rating of a grade, or school, or system to bring together somehow the results of the several tests into a single table, or graph. The citizen who votes the support of the schools is particularly interested in this final, simple form of the results, while the experts only—the teachers, supervisors, and administrators—who should have the time to study details relative to their respective duties, are interested in the smaller diagnostic features of the report.

Recently the writer found himself confronted with the necessity of devising a satisfactory method of condensing reports on the Cleveland Survey Arithmetic Tests, or of adopting tests inferior for the purpose in view. The usual situation was complicated by the fact that a large proportion of the children to be tested were negroes, whose status was to be separately shown. The school system to be surveyed was that of Paducah, Kentucky, where at the time of the survey 3,303 pupils were in attendance. Limited funds compelled us to make the report on educational measurements comparatively short.

To meet these practical demands of limited space and yet to present numerically all the medians (absolute scores) in relation to norms of various grades, we decided to use the form shown in Table II. This form of report could be used to show conditions by buildings as well as for the system as a whole.

In the present paper, however, we have omitted the scores of the several schools, and have shown only the medians of all the scores of whites and of negroes of each grade. In a small system it is possible also to include in the general table the medians of the several schools, as is illustrated in Table III for grade IVA; but in most surveys a table of this sort would become so bulky as to prevent a careful study of the more generally significant aspects of the data. In such cases each yearly grade could probably be shown in a separate table similar to Table III. Even then, however, a general summary such as that shown in Table II, would also be advisable, both to bring the whole situation within easy grasp and also to facilitate the construction and comprehension

TABLE I. RATIO OF MEDIANS AT THE DEMONSTRATION SCHOOL TO THOSE AT CLEVELAND AND GRAND RAPIDS

Grade	No. of Pupils	Comparison City	Tests[a]															
			A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Av.
VIIIB	34	Cleveland	113	138	129	98	103	108	91	112	76	70	64	77	66	39	36	88
		Grand Rapids	105	118	127	108	107	108	90	111	88	70	82	61	60	50	33	88
VIIB	32	Cleveland	101	95	114	101	34	93	85	39	75	82	79	94	67	50	49	77
		Grand Rapids	102	96	114	114	69	82	82	33	79	75	75	70	67	51	50	77
VIB	31	Cleveland	90	85	72	78	44	67	73	36	32	73	71	72	53	59	64	65
		Grand Rapids	85	111	77	94	50	70	75	32	43	73	111	61	47	91	57	72
VB	22	Cleveland	96	84	84	130	61	75	77	60	13	50	29	40	42	.....	.....	65
		Grand Rapids	106	98	84	164	77	83	89	.....	71	58	67	43	50	.....	.....	83
IVB	32	Cleveland	95	84	83	60	56	61	64	.....	.....	31	75	58	40	.....	.....	63
		Grand Rapids	125	120	140	108	73	107	113	.....	.....	.....	.....	.....	.....	.....	.....	112
Average		Both	103	103	102	106	67	86	84	60	60	65	65	64	54	57	48	79

[a] Explanation of tests: A is addition, two digits; B, subtraction, one operation; C, multiplication, two digits; D, simple division, one-digit answers; E, addition of single columns of five digits in each column; F, subtraction with borrowing, simple; G, multiplication, a number of four digits by one of one; H, fractions with common denominators, addition and subtraction; I, division, one-digit numbers into five-digit numbers; J, addition of single columns of thirteen digits in each; K, simple long division, divisor of two-digit numbers; L, multiplication of four-digit by two-digit numbers; M, addition of five four-digit numbers; N, long division by two-digit numbers, all quotients coming out even; O, simple fractions, the four operations.

TABLE II. MEDIAN SCORES AT PADUCAH IN THE CLEVELAND ARITHMETIC TESTS

Grade	No. of Pupils	Tests[a]														
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
VIIIA Whites	51	22.5	19.1	16.8	18.8	4.9	7.9	5.0	2.4	4.1	4.1	7.4	3.5	4.0	1.4	1.2
Negroes	21	24.2	17.8	15.7	14.5	4.8	7.3	3.9	1.0	1.6	2.9	3.9	2.3	3.1	0.7	2.5
St. Louis		*32.2*	*28.3*	*21.9*	*25.7*	*8.4*	*11.3*	*7.8*	*12.0*	*5.8*	*5.8*	*11.7*	*5.3*	*5.3*	*2.7*	*6.6*
Grand Rapids		*30.3*	*25.5*	*20.7*	*23.0*	*8.1*	*11.0*	*6.8*	*8.8*	*4.7*	*6.5*	*10.3*	*4.9*	*5.7*	*2.3*	*4.6*
VIIIB Whites	81	27.1	18.6	15.8	16.3	5.3	6.7	5.0	3.7	2.4	3.4	6.0	2.8	3.3	1.4	1.9
Negroes	9	20.8	21.3	13.5	16.5	6.2	8.5	3.3	8.2	1.5	3.5	7.7	2.3	4.5	0.9	2.4
St. Louis		*32.3*	*26.7*	*20.7*	*23.8*	*8.0*	*10.1*	*7.4*	*10.8*	*5.4*	*5.4*	*10.3*	*5.2*	*5.2*	*2.6*	*6.1*
Grand Rapids		*29.5*	*22.8*	*19.3*	*20.5*	*7.8*	*10.3*	*6.7*	*8.6*	*4.0*	*5.7*	*9.7*	*4.9*	*5.7*	*2.0*	*5.5*
VIIA Whites	96	24.3	17.5	15.7	17.7	5.0	6.4	4.9	3.0	2.1	3.5	6.8	3.1	2.9	1.4	1.7
Negroes	23	23.5	13.7	14.5	11.5	4.1	4.7	4.2	5.4	0.8	2.6	3.1	1.3	2.6	0.7	0.6
St. Louis		*28.4*	*24.2*	*19.8*	*22.3*	*7.4*	*9.6*	*6.9*	*9.7*	*5.0*	*5.3*	*9.7*	*4.7*	*4.9*	*2.0*	*5.6*
Grand Rapids		*27.3*	*20.7*	*18.8*	*19.7*	*7.2*	*9.6*	*6.1*	*7.8*	*4.1*	*5.3*	*8.8*	*4.5*	*5.0*	*1.8*	*4.6*
VIIB Whites	104	21.4	17.7	13.7	15.3	4.8	5.7	4.3	4.0	1.6	3.3	4.6	2.0	2.7	0.8	1.0
St. Louis		*27.8*	*22.8*	*18.9*	*21.3*	*6.6*	*8.5*	*6.4*	*9.5*	*4.5*	*5.2*	*8.3*	*4.6*	*4.5*	*2.0*	*4.8*
Grand Rapids		*26.5*	*21.3*	*17.7*	*18.4*	*7.2*	*9.3*	*6.1*	*9.0*	*3.8*	*5.4*	*7.5*	*4.3*	*4.9*	*1.7*	*3.9*
VIA Whites	97	19.8	15.7	13.3	12.7	5.1	5.1	4.1	4.3	1.3	3.2	3.8	1.9	3.3	0.8	0.9
Negroes	39	18.3	16.3	16.5	14.3	4.7	5.5	4.1	2.4	1.3	2.9	3.8	1.6	2.2	0.7	0.8
St. Louis		*26.4*	*20.6*	*18.3*	*20.5*	*7.1*	*8.3*	*6.2*	*8.1*	*4.1*	*5.1*	*7.4*	*4.1*	*4.4*	*1.8*	*4.1*
Grand Rapids		*25.0*	*19.1*	*17.0*	*16.9*	*6.6*	*8.0*	*5.6*	*6.5*	*3.0*	*4.5*	*6.5*	*3.6*	*4.5*	*1.4*	*3.6*
VIB Whites	90	16.8	14.9	13.0	12.3	4.5	4.1	3.8	3.2	0.8	3.0	3.0	1.6	2.8	0.8	1.1
St. Louis		*26.3*	*20.3*	*18.2*	*19.3*	*6.9*	*8.0*	*5.9*	*8.0*	*3.9*	*5.0*	*6.9*	*4.3*	*4.2*	*1.6*	*3.6*
Grand Rapids		*22.8*	*16.8*	*15.5*	*15.5*	*6.0*	*7.1*	*5.3*	*6.2*	*2.3*	*4.1*	*5.4*	*3.3*	*4.3*	*1.1*	*3.5*

[a] For description of tests see Table I.

Grade	No. of Pupils	Tests														
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
VA Whites	95	14.8	14.1	9.1	12.2	4.8	4.7	3.2	4.3	0.7	2.8	2.4	1.5	2.2	0.7	1.5
Negroes	35	13.9	12.9	13.6	12.5	3.6	4.5	3.3	2.9	0.7	1.9	2.7	1.2	1.3	0.7	0.6
St. Louis		*22.5*	*20.0*	*16.7*	*17.8*	*6.1*	*7.4*	*5.6*	*6.5*	*3.2*	*4.3*	*5.8*	*3.4*	*3.7*	*1.3*	*3.3*
Grand Rapids		*21.5*	*15.9*	*14.0*	*14.3*	*5.4*	*6.5*	*4.9*	*6.3*	*1.4*	*3.7*	*4.3*	*2.9*	*3.6*	*0.8*	.....
VB Whites	72	12.8	8.0	8.4	5.8	3.3	2.0	2.4	0.8	0.5	2.5	0.8	0.9	1.8	0.5	0.6
Negroes	21	17.8	8.5	7.5	1.9	4.2	2.3	1.9	1.5	0.5	2.5	1.1	0.7	1.5	0.6	.....
St. Louis		*22.5*	*18.0*	*16.9*	*18.4*	*6.0*	*6.4*	*5.5*	*4.8*	*3.0*	*4.1*	*5.0*	*3.1*	*3.4*	*1.3*	*3.3*
Grand Rapids		*20.3*	*14.7*	*13.7*	*12.5*	*5.2*	*6.0*	*4.5*	.....	*1.3*	*3.4*	*3.0*	*2.3*	*3.0*	*0.7*	.....
IVA Whites	102	13.7	10.2	7.8	7.1	3.7	3.6	2.4	1.8	0.6	2.3	1.0	0.8	1.3	0.5	0.6
Negroes	36	13.1	10.7	7.5	6.3	4.0	5.1	2.1	2.6	0.5	3.0	0.9	0.6	1.4	0.5	0.6
St. Louis		*21.3*	*17.0*	*15.4*	*16.3*	*5.4*	*6.0*	*5.1*	*6.8*	*2.0*	*3.9*	*4.0*	*2.9*	*3.3*	*1.1*	*2.5*
Grand Rapids		*16.4*	*12.1*	*11.3*	*10.4*	*4.6*	*4.1*	*3.3*	.....	*0.9*	*2.8*	.....	.....	*2.3*	.....	.....
IVB Whites	124	12.9	5.3	5.0	3.5	2.8	1.0	1.4	2.0	0.6	1.5	.....	0.6	0.9	.....	0.5
Negroes	24	13.0	5.0	4.0	3.5	2.3	2.0	1.4	1.0	0.5	0.9	0.7	0.7	0.8	0.6	0.5
St. Louis		*19.8*	*17.1*	*16.7*	*15.8*	*5.7*	*5.6*	*4.9*	*7.8*	*2.0*	*3.8*	*3.3*	*2.5*	*2.9*	*0.8*	.....
Grand Rapids		*13.6*	*9.1*	*7.1*	*6.9*	*4.1*	*2.8*	*2.2*	.....	*0.7*	.....	.....	.....	.....	.....	.....
IIIA Whites	101	11.4	7.8	8.0	4.1	2.9	0.9	1.6	0.5	.....	0.7	.....	0.6	0.6	.....	.....
Negroes	39	5.5	4.8	3.9	2.8	2.1	0.8	1.2	0.7	.....	0.7	.....	0.5	0.6	.....	.....
St. Louis		*18.3*	*12.2*	*10.5*	*12.2*	*4.8*	*3.5*	*3.5*	*3.8*	*1.4*	*2.9*	.....	.....	*2.1*	.....	.....
Grand Rapids		*13.4*	*8.4*	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
IIIB Whites	105	8.9	4.9	2.1	1.6	2.5	0.6	0.7	0.8	.....	0.5	.....	.....	.....	.....	.....
Negroes	31	7.5	2.1	1.4	0.7	1.6	0.6	0.6	0.5	.....	0.6	.....	.....	.....	.....	.....
St. Louis		*14.6*	*9.9*	*7.6*	*9.0*	*3.8*	*2.3*	*2.7*	*0.7*	*1.1*	*1.6*	.....	.....	*0.5*	.....	.....
Grand Rapids		*11.8*	*6.3*	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

TABLE III. ILLUSTRATING METHOD OF SHOWING MEDIANS OF SEPARATE SCHOOLS FOR EACH GRADE—CLEVELAND ARITHMETIC TESTS. DATA FOR GRADE IVA

School	No. of Pupils	Tests[a]														
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
WHITES	102	13.7	10.2	7.8	7.1	3.7	3.6	2.4	1.8	0.6	2.3	1.0	0.8	1.3	0.5	0.6
Washington	24	13.6	10.8	8.3	7.8	3.6	4.4	2.8	4.4	0.5	2.4	1.4	0.8	1.6	0.5	0.5
Whittier	12	13.3	8.3	8.7	9.0	3.3	3.5	1.8	3.7	0.7	2.3	2.0	1.0	1.2	0.6	0.6
Franklin	26	14.0	10.0	9.0	3.5	4.7	2.5	1.7	0.7	0.6	1.4	0.7	0.8	1.3	0.5	.....
McKinley	26	11.7	7.0	6.2	5.0	3.6	3.3	2.5	3.0	.....	1.6	1.1	0.7	0.9	0.5	0.7
Lee	14	19.5	15.5	7.5	13.0	3.0	5.0	2.8	2.7	1.0	2.0	1.0	0.8	2.3	0.6	.....
NEGROES	36	13.1	10.7	7.5	6.3	4.0	5.1	2.1	2.6	0.5	3.0	0.9	0.6	1.4	0.5	0.6
Lincoln	21	14.5	11.3	7.9	7.3	5.2	5.3	2.3	2.2	.....	3.8	1.6	0.7	1.9	0.6	0.7
Garfield	15	11.6	8.5	3.5	4.3	2.9	5.1	1.7	3.5	0.5	1.7	.....	0.6	1.2	.....	0.5
NORMS																
St. Louis	.....	21.3	17.0	15.4	16.3	5.4	6.0	5.1	6.8	2.0	3.9	4.0	2.9	3.3	1.1	2.5
Grand Rapids	.....	16.4	12.1	11.3	10.4	4.6	4.1	3.3	.....	0.9	2.8	.....	.....	2.3	.....	.....

[a] For description of tests see Table I.

of a subsequent table in which retardation or acceleration would be shown in terms of the grade units which we have advised above.

Such a table was utilized in the Paducah survey. It is here presented as Table IV. It brings out emphatically in small space the median retardation of each grade, and records negroes and whites separately.

A table of this sort can be constructed in a comparatively short time by an intelligent assistant. The method of doing so and the significance of the table can best be made clear by a few concrete illustrations. According to Table II the score of the white children in grade VIIIA on test A is 22.5. This is below either of the standards given for the VIIIA grade. Glancing down the same column at the standards of lower grades, we note that this score exactly equals the St. Louis standard for grade VA, and that it is approximately equal to the Grand Rapids standard for grade VIB (22.8). These grades are therefore put down in Table IV under test A, opposite grade "VIIIA, Whites" in the lines for St. Louis and Grand Rapids, respectively. The standing of the negroes is indicated in a similar manner. Again, in test B, Paducah grade VIIIA, whites, equals approximately St. Louis and Grand Rapids grades VA and VIA, respectively; and so on. The nearest half-grade in the norms is always chosen. It will be noted that in the case of test H, Paducah is especially low; and that on practically every test the negroes are below the whites. At the right of the table we indicate in terms of years the average distance that the respective Padacuh grades were found to fall short of the norms considered. This rating is based on the assumption that the grades indicated under each test are correctly entered. But only the nearest half-grade is given; accordingly, there are slight errors, sometimes in one way, sometimes in the other. It is assumed, however, that these errors will on the whole cancel one another when averaged for the fifteen different tests, since the chances of error in the two directions are practically equal.

It was, of course, impossible thus to rate the lower grades, whose scores were below any available norms. But this is not important as deficiencies in the lower grades are, as a rule, reflected in the higher grades if they are characteristic of the whole system; and it is the upper-grade children that must soon come into competition with children from other systems. A deficiency

TABLE IV. COMPARISON OF PADUCAH SCHOOLS WITH ST. LOUIS AND GRAND RAPIDS STANDARDS

Paducah grade	Comparison City	Tests (for description of tests see Table I)																Paducah is behind (years)
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	All	
8A, Whites	St. Louis	5A	5A	5A	6B	3A	6B	4A	3A	6A	5B	6A	5A	6B	5A	4A	5A	3.17
8A, Whites	Grand Rapids	6B	6A	6A	7B	4A	6A	5A	5A	7A	6B	7B	6A	6B	6A	6B	6A	2.23
8A, Negroes	St. Louis	6B	5B	4A	4B	3A	5A	3A	3B	3A	3A	4A	4B	4A	4B	4A	4A	4.23
8A, Negroes	Grand Rapids	6A	6A	6B	5A	4A	6B	4A	5B	5A	4A	5A	5B	5B	5B	6B	5A	3.10
8B, Whites	St. Louis	7B	5B	4A	4A	4A	5B	4A	3A	4A	4B	5A	4A	4A	5A	4A	4A	3.43
8B, Whites	Grand Rapids	7A	6A	6B	6A	5A	5A	5A	6B	6B	5B	6A	5A	5A	6A	6B	6B	2.03
8B, Negroes	St. Louis	4A	6A	4A	4A	5A	7B	3A	6A	3A	4B	6A	4B	7B	4B	4A	5B	3.03
8B, Negroes	Grand Rapids	5B	7A	5B	6A	6B	6A	4A	8B	5A	5B	7B	5B	6A	5A	6B	6B	2.03
7A, Whites	St. Louis	6B	5B	4A	5A	3A	5B	4B	3A	4A	4B	6B	5B	4B	5A	4A	4A	2.80
7A, Whites	Grand Rapids	6A	6A	6B	7B	5B	5A	5A	5A	6B	5B	6A	6B	5B	6A	6B	6B	1.60
7A, Negroes	St. Louis	5A	3A	4B	3A	3B	4B	4B	4B	3B	3A	4B	4B	4B	4B	4B	4B	3.60
7A, Negroes	Grand Rapids	6B	5B	5A	5B	4B	4A	5B	5A	4A	4A	5B	5B	5B	5B	5A	5B	2.50
7B, Whites	St. Louis	4A	5B	4A	4B	3A	4B	4B	3A	3A	4B	5B	4B	4B	4B	4A	4B	2.93
7B, Whites	Grand Rapids	5A	6B	5B	6B	4A	5B	5B	5A	5A	4A	5A	5B	5B	5A	6B	5A	1.70
6A, Whites	St. Louis	4B	4B	3A	3A	4B	4B	3A	3A	3A	3A	4A	4B	4A	4B	4A	4B	2.60
6A, Whites	Grand Rapids	5B	5A	5B	5B	5B	5B	5B	5A	5B	5B	5A	5B	5A	5A	6B	5A	1.27
6A, Negroes	St. Louis	3A	4A	4B	4B	3A	4B	3A	3A	3A	3A	4A	4B	3A	4B	4B	4B	2.37
6A, Negroes	Grand Rapids	5B	6B	6A	5A	4A	5B	5B	5A	5B	4A	5A	5B	4A	5B	5A	5A	1.20
6B, Whites	St. Louis	3A	4B	3A	3A	3A	3A	3A	3A	3B	3A	4B	4B	4B	4B	4A	3A	2.30
6B, Whites	Grand Rapids	4A	5B	5B	5B	4A	4A	4A	5A	4A	4A	5B	5B	5B	5A	6B	5B	1.07
5A, Whites	St. Louis	3B	3A	3A	3A	3A	4B	3A	3A	3B	3A	4B	4B	3A	4B	4A	3A	1.87
5A, Whites	Grand Rapids	4B	5B	4B	5B	4A	4A	4A	5A	4B	4A	5B	5B	4A	5B	6B	5B	0.77
5A, Negroes	St. Louis	3B	3A	4B	3A	3B	3A	3A	3A	3B	3B	4B	4B	3A	4B	4B	3A	1.97
5A, Negroes	Grand Rapids	4B	4A	5B	5B	4B	4A	4A	5A	4B	4A	5B	5B	4A	5B	6B	5B	0.77
5B, Whites	St. Louis	.....	.....	3B	.....	.....	.....	.....	3B	.....	3A	.....	.....	3A	.....	.....	4B	1.75
5B, Whites	Grand Rapids	3A	3A	4B	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	3A	1.33
5B, Negroes	St. Louis	3A	.....	3B	3B	3A	3B	.....	3B	.....	3A	.....	.....	3A	.....	.....	3B	1.77
5B, Negroes	Grand Rapids	4A	3A	4B	.....	4B	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	4B	1.00

Table reads: at Paducah the 8A whites attained scores approximately equal to those attained at St. Louis by the 5A grade for Test A, by 5A grade for test B, by the 5A grade for test C, ...and by the 5A grade for all tests; these children were according to St. Louis standards 3.17 years behind; etc.

in a certain subject in the lower grades may be due, in fact, to a general emphasis on other lines of work at first, and may be regularly compensated for later.

On the other hand, systems in which the scores are superior to these norms used cannot be accurately rated in the eighth grade by the method exhibited in Table IV; for the comparison norms do not extend above this grade. But this is not a serious defect in the method, for emphasis on any subject may safely be relaxed when the grade norms are reached. For instance, an VIIIA grade exceeding by an appreciable amount the St. Louis median of 32.2 in test A might reasonably be excused from further drill on additional work of this sort; and since there is no "grade" significance to progress beyond this standard, it is sufficient to know merely that the norm has been surpassed. In a school system where such acceleration exists its approximate amount can be shown by noting in the table which of the grades below VIIIA first reached the norm of this grade.

TABLE V. DEFICIENCY OF PADUCAH SCHOOLS IN TERMS OF YEARS WHEN COMPARED WITH STANDARDS OF ST. LOUIS AND GRAND RAPIDS

Grades	St. Louis		Grand Rapids	
	Whites Years behind	Negroes Years behind	Whites Years behind	Negroes Years behind
VIIIA	3.17	4.23	2.23	3.10
VIIIB	3.43	3.03	2.03	2.03
VIIA	2.80	3.60	1.60	2.50
VIIB	2.93	.....	1.70	.....
VIA	2.60	2.37	1.27	1.20
VIB	2.30	.....	1.07	.....
VA	1.87	1.97	0.77	0.77
VB	1.75	1.77	1.33	1.00
Average	2.61	2.83	1.50	1.77

Our next step may be to bring together the general results of the data shown in Table IV. This is done in Table V, which shows a progressive retardation of Paducah children with increasingly higher grades. As between whites and negroes the dif-

| SB, Negroes.... | Grand Rapids.. | 4A | 3A | 4B | ..... | 4B | ..... | ..... | .....

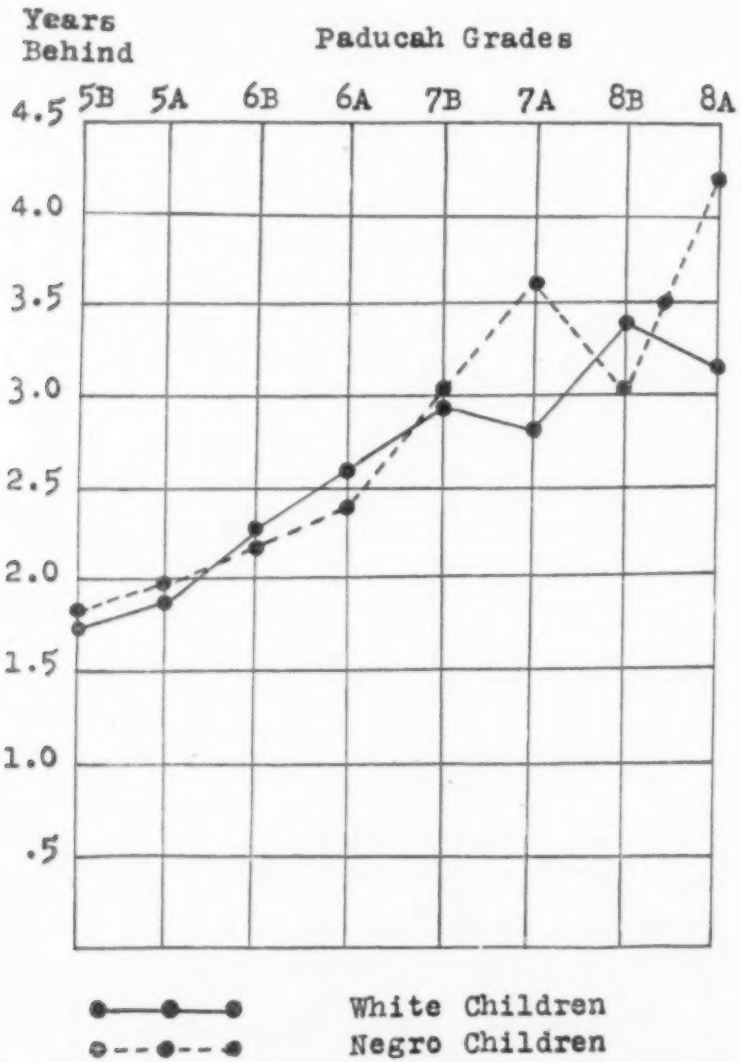


FIGURE 1. SHOWING PROGRESSIVE RETARDATION OF PADUCAH SCHOOL CHILDREN IN ARITHMETIC AS DETERMINED BY THE CLEVELAND TESTS. RETARDATION IS GIVEN IN UNITS OF A YEAR'S PROGRESS. ST. LOUIS STANDARDS USED

ferences in favor of the former are greater in the higher grades. In grades V and VI combined the two races are on practically equal terms. In grades VII and VIII combined the negroes fall short of the whites by 0.49 of a year according to the St. Louis standards and 0.65 of a year according to the Grand Rapids standards.

Figure 1 shows graphically part of the comparisons given in Table V. It is arranged in the ascending order of the grades to show that the retardation gradually and regularly increases with years of training in the system.

It is to be understood that Tables IV and V, and Figure 1, are constructed on the assumption that the tests all have equal weight in determining the general status of the schools in arithmetic; that is, that the extent of retardation in each grade in terms of yearly progress may be found simply by averaging the deviations from the norms in each of the fifteen different tests. No special weighting was given any one of these tests. While some of them are, of course, harder than others—on some tests, indeed, the children in the lower grades did not make any score at all—it has been thought fair to find grades in the other school systems, making scores equal in each test to those of the several grades at Paducah, and to make our comparisons on this basis. It seems fair, in other words, to call a grade in one school system equal to a lower grade in another if the two grades actually do, under uniform conditions, the *same* amount and quality of work. In the fundamentals of arithmetic, we may then conclude, the schools here used to illustrate our methods are seriously behind the standards of representative American cities, so far as rapidity and accuracy in the fundamentals are concerned. This deficiency is clearly shown to be general throughout all the schools and grades. In grades III, IV, and VB the extent of the deficiency cannot be exactly measured, because these grades fall below available standards in some of the tests used. The general status of the schools is clearly indicated in Figure 1, which shows that the retardation is greatest for the higher grades and that it decreases as we go down the grades. This is as we should expect, if the causes of the deficiency are general in the schools; for these causes will have operated more years in the higher than in the lower grades. If the native ability of the children in this system is on

the average comparable to that of the children in the cities used for comparison—and for this assumption we have supporting evidence from a mental test—the difference in median scores should be negligible at the time of beginning arithmetic work in the schools. For simplicity, Figure 1 represents graphically the comparison with only one other school system, but it is evident from Table V that the school studied is seriously behind reasonable norms.[4]

The suggested method of handling the data of the fifteen Cleveland tests shows in a simple, convincing manner the general status of the schools as regards the fundamentals of arithmetic; and it is also capable of showing in any degree of detail desired by professional educators the work of different grades and schools in any phase of the processes emphasized by these tests. One grade in a certain school may, for instance, be found to be comparatively normal in the very simplest forms of the fundamental processes, and yet seriously deficient on such phases of these processes as adding long columns, subtraction with borrowing, estimating the probable number of times one number is contained in another, etc. It is certainly important in surveys that just such diagnostic features of school work be brought out in a way available for the experts who are to correct the defects shown in the more general graphs for the non-technical reader. If such necessary data are furnished in the report, the professional educators concerned can work out for their special purposes and needs such additional graphs and tables as they desire.

[4] It is but fair to add here that the Paducah schools, used in our illustration, were tested at a time when the disturbances from influenza were to be reckoned with, but there is no satisfactory evidence that the status of the schools would otherwise have been noticeably different from that shown in the report. The general spirit of the directors of the schools was, fortunately, such as to remedy the defects as soon as possible. Important steps to this end were taken as a result of the survey.

Editorials

PRACTICAL VERSUS PURE RESEARCH

Pure research is a term used to describe the investigation of problems whose selection is based primarily on the individual interest of the investigator and whose solutions need have no value as direct influences upon human conduct. An investigator in the field of pure research has to account neither for the source of his problem nor for his solution—except that the solution must be arrived at in accordance with the rules of the game in his particular field. His justification is adequately based upon the fact that the human mind is permitted to expend its energies upon any undertaking which is conceived and carried on in the spirit of pure research. That the results obtained may later be turned to use by other agencies in influencing conduct does not detract from the spirit of such research. Many of the principles of electricity which are now used in the operation of public utilities and domestic flatirons were discovered by scientists who did not foresee their practical applications. Or, if dimly seeing, they investigated merely for the love of discovery.

Practical research, on the other hand, is a term used to describe the investigation of problems which spring from the difficulties of human conduct and whose solutions are evaluated in terms of their efficiency in improving conduct. The beginning and the end of practical research reside in conduct. It begins with faulty practice and ends with the improvement of practice. Such problems as those of educational measurements rise as practical research problems out of the practices of the schoolroom when teachers are unable to determine in objective terms the ability of their pupils and the effectiveness of their methods, or when administrators are unable to judge of the relative efficiency of their teachers. The value of educational measures is judged absolutely by their effect upon school practices as shown in accurate measures, improved methods, and exact evaluation.

Bureaus of educational research are agencies of practical rather than of pure research. The members of their staffs are not justified in selecting any problem through mere individual preference. The problems must be found in the operations of the school, and the solutions are neither complete nor valid until they have affected school practice to its advantage.

In both of these respects, the bureaus are frequently at fault. Their problems are sometimes chosen without regard to institutional importance and the solutions are not carried to such lengths that they can be used in practice. At other times, and much more frequently, their problems are well chosen but the solutions are not perfected with the degree of adaptation necessary for use.

For instance, during the last five years, educational tests have been widely given by bureaus. It has been discovered and shown in numberless tables that some children attain greater efficiency than others and that some classes are superior to others. The amount of labor involved in giving and tabulating the tests has been enormous and frequently the distress of the teachers has been acute. In many cases the morale of the instructional staff has been injured and the reputation of the bureaus as service agencies has been ruined. The teachers feel the true facts of the case. Such bureaus have not been constructive because they have not improved practice; they have taken only the first step and found the problem. When they have put their hands to the plow, they have stopped in the furrow before the task was complete. Practice is no better than before except in so far as the urge of fear has driven the teachers to improve methods whose inferiority is evidenced by the poor showing of their students. The bureau's solution is not complete until it has both pointed out the weakness of the methods and collected new methods or disseminated old ones and has seen that the technic of using them is known. When practice has been adequately improved, then and not until then is the research bureau's work complete.

Or, to take a very simple case, a child is sent to the bureau for mental examination because his work is not satisfactory. The tests being given, an unfavorable report may be made to the principal. But the bureau's task is not complete if the child is returned to the same environment and nothing is done. Practice has not been improved. There yet remains the question of what

to do with the case. This may lead to a long campaign for special rooms for defectives, or to the quick decision to place this defective in a school for his kind or to put him in a class more nearly fitted to his capacities. But in any case the urge to see the task carried to the end must not be weak; it is strong enough in good bureaus to keep pressure applied persistently and diplomatically until something is done.

This means that the members of the staff cannot be mere specialists who are satisfied with reaching an abstract solution. They must have intimate knowledge of the workings of the school system so that their recommendations can be applied without material modifications. For in the event of failure to do something, the administration cannot be blamed except on the proven grounds of laziness. The bureau will be wise if it assumes that the fault is its own because its solutions are not practical and as such do not fit into the educational situation.

So important is the difference between the standpoint of pure and practical research and so easy is the acceptance of incomplete practical solutions that a statement of certain rules of procedure in practical research is undoubtedly worth while. Of these five may be mentioned.

First, *study the operations of the school to discover the points at which they are failing to function efficiently*. This study may be in the form of the occasional formal survey which reveals difficulties as they appear to imported experts. Or it may be the sort of study which is carried on incidentally by the superintendent and supervisors who can from long experience place their fingers upon the weak points of the system. Or it may be the result of investigations carried on by the bureau itself. But in any case the bureau should have at its command a picture of the system at work and a list of the points at which it fails to function with efficiency.

Second, *select and define a problem for investigation*. In the selection of the problem three factors enter. First, other conditions being neutral, the most important problems of the system should be selected. Second, the resources of the bureau are, also, a determining factor; for large investigations cannot be undertaken with an inadequate personnel. Third, the time element must likewise have weight, because confidence in the work of the bureau is dependent upon fairly frequent solutions of some prob-

lems while others of greater magnitude are pending. To these may possibly be added the individual preferences of the staff as a fourth factor, but this is not deserving of serious consideration except in the selection from among several problems, each equally appropriate from the more worthy points of view. The definition of the problem follows as a matter of course as part of the process of investigation.

As an illustration of the rules, I shall describe the development of an investigation with which I am rather intimately acquainted. In 1915, Superintendent Cammack of Kansas City arranged for the selection of language and grammar as a problem for study. He did so because he recognized from his long experience that the language situation was serious in a large and cosmopolitan community such as Kansas City. Moreover, the time element was not prohibitive, and the investigation was not larger than could easily be conducted by one investigator with a few assistants. The definition of the problem resulted in the decision to find what grammar would be taught if it were based on the errors of school children with the aim of teaching only those rules of the subject which were broken and the definitions necessary to understand the rules.

The third rule follows naturally: *using the defined problem as the central element, proceed to its solution.* This is essentially a laboratory process. The problem is temporarily taken out of its setting in the practical operations of the schoolroom, the literature is canvassed, subordinate studies are made, experts are consulted, and a solution is reached which it is hoped will, when applied to the schools, improve instructional practice. While this procedure is being carried on, constant regard for practice is necessary in order to check up the trend of the solution and the form it will finally take. If the solution will not work in practice it must be discarded for another that will. As a result of this constant consideration, the solution finally accepted may be quite different in form and principle from that which was originally contemplated. Preliminary studies must often be tried out in a small way to discover defects before proceeding to a comprehensive attack. But finally the solution which satisfies the demand of the defined problem will be obtained.

Reverting to our illustration, we observe that many of these processes were followed. The oral and written errors of the

children were collected by the teachers. The oral errors were classified. Samplings of written material were read and the mistakes tabulated. Errors in technic were corrected. Definitions necessary to understand the rules were determined. The material was systematically arranged and comparison of this material with that in the adopted textbook of the city was made. Finally, the findings were printed in a bulletin and given to the teachers.

At this point it seemed at the time that our job was completed. We had published. But school practice had not been improved, and additional steps were therefore necessary.

This leads to the fourth rule. *Install the solution in the practice of the school.* This is absolutely essential if practice is to be bettered by the work already done. Publications, charts, statistics, and exhibits do not, of themselves, mark the end of an investigation. They would be an adequate justification of pure research, but in practical research the final test is found only in the improvement of practice.

Installation is the Waterloo of many investigators. While they are solving the problem in the laboratory, they are dependent only upon their own efforts; but when they begin to install, they must reckon with the peculiarities of people. These people may lack the training necessary to use a solution too highly technical. Deep-seated prejudices may be encountered. The program of studies, the textbooks in use, the available supplies, and the crowded character of the curriculum may all have to be taken into account. The hasty investigator may grow irritated at these conditions and blame the system, the board, the superintendent, and the teachers. The wise investigator, however, considers all these factors and presents a solution which will work rather than one which suits him. Or, if he has failed to foresee all the difficulties, he is cheerfully willing to make the adaptations necessary to produce a practical solution.

This does not mean that at times the bureau should not stand for its own solutions against the opinions of the rest of the system. Such situations constantly arise. But it does emphatically mean that the bureau must sell its idea to the instructional and administrative staffs if it hopes to put them into operation. Modification of non-essentials must usually be made if the essentials are pre-

served. Fortunate is the man who finds his methods are accepted and used without modifications. Cooperation is as necessary as a good solution if effective use is to be made of the findings. Since teachers and administrators must apply the solution in order to improve practice, a practical solution is valueless if it is not used. Thus, we find conferences, committees, and educational campaigns to be necessary factors in installation. While this sort of work is distasteful to many investigators, it is an absolutely necessary part of the work of a bureau; and no one incapable of, or uninterested in, getting the cooperation of the personnel of the schools is justified in being on the bureau staff in any other than a clerical capacity.

In the case of the grammar study, we thought at first that with the publication of the report our work was completed. But it soon became evident that the teachers merely read the report without making use of it. A number of factors at once emerged. Not all the staff of seventh- and eighth-grade teachers accepted our basis for the selection of the content of grammar. The textbook in use, which could not be changed, used another basis. If the report was adopted and certain parts of the text were left out, the exercises which followed the omitted portions might occasionally depend upon them.

As a result of these conditions a committee composed of high-school teachers and grammar-grade teachers, selected by the teaching force, met frequently and discussed the proposition with great frankness. As a result the report was adopted, the textbook was thoroughly edited and portions omitted, together with later exercises dependent upon these. This was then referred to the grammar-grade teachers, and the report was adopted by all. The changes made necessary in the text were printed and distributed.

Obviously, a much better solution would have been the publication of a grammar text which carried out the plan of the report; but it was necessary for us to reach not the best solution so much as one which in view of all the conditions would work.

The installation advances the solution one step but it does not end the matter. Although practice may for a time be improved, the improvement may turn out to be merely temporary. Therefore, the fifth rule is necessary: *provide or utilize machinery*

for the supervision of the solutions. The bureau is concerned with the permanent improvement of practice, and this can be obtained only by supervision. Mere inspection is frequently sufficient, but in many cases the teachers must be instructed in the best methods of operating the solution.

This means that while the bureau does not take on the duties of a supervision, it must see that supervision is maintained and keep close watch on the operation of the solution until it has become an integral part of school practice. Its duty is not so much that of inspection as that of constructive assistance in ironing out difficulties of adaptation to obstructive conditions or to conditions not foreseen until the operations were under way.

Reverting to the grammar study, this fifth step was not taken in any formal way. The process of supervision is long continued. The cost and inconvenience of traveling a long distance made it difficult to follow the process of supervision. But while nothing formal was done, the subject received the same supervision from the supervisors as did the other subjects of the curriculum.

It is a detail of administration as to whether the fourth and fifth steps should be carried on by the bureau or by some other agency in the system. In some cases the superintendent or the supervisor would be better able to install; in others, the bureau would be the logical agency. But in either case, the job is not completed until practice is permanently improved. Therefore, bureaus of educational research owe it to their own reputations for efficiency and service to see that the task is not forsaken until their results have become a part of the regular practice of the school.

W. W. C.

EXCUSE ME!

In Dr. Chapman's article on "Group Intelligence Tests Without Prepared Blanks" in our December, 1920, number, the footnote was ambiguous. It should have read, "Copies of this test may be secured in any number at twenty-five cents PER COPY, on application to the author at the Department of Education, Yale University." The words "per copy" were omitted in error. Some literal minded people have been ordering large quantities of the test, believing that they would receive the entire shipment for twenty-five cents. Each copy of the test is twenty-five cents—and worth the money.

EDITOR.

Reviews and Abstracts

E. H. CAMERON, *Editor*

EVENDEN, E. S. *Teachers' salaries and salary schedules in the United States, 1918-19.* Washington: National Education Association, No. 6 of series of Commission on the Emergency in Education, 1919. 169 pp.

A review of this report now necessarily differs greatly from a review at the date of publication, October, 1919. The writer was not asked to undertake the review until the fall of 1920 and other duties have prevented the writing until January, 1921. But so valuable is the report that even at this date a review is thoroughly worth while.

A book review for readers of the JOURNAL OF EDUCATIONAL RESEARCH should in the writer's judgment answer these test questions which are given in order of importance to the reader:

1. What are the essential results of the investigation that are of importance to readers of the Journal?
2. What classes of readers should read the book or purchase it for their school or personal libraries?
3. Did the educational problem treated merit the time, energy, and ability of the investigator expended upon it?
4. How nearly did the work, as to the territory covered and methods employed, solve the problem for which the investigations were undertaken?
5. What methods of investigation, presentation of results and circulation of results, would be useful in similar investigations?
6. What further investigations of importance in the field covered by this investigation does it leave to be done?

1. Many readers of the Journal have already seen the report and so have formulated answers for themselves to the first two questions. Furthermore, they have in many cases seen summaries or extracts from it in various national and local educational periodicals as well as in the lay press. A considerable number have doubtless assisted in making these summaries or have used the book in conducting local salary studies. But for those who are not thus familiar with it, for students of education who are preparing for administrative positions, and for clearness in this discussion, a birdseye view is desirable.

The report was planned by the Commission on the Emergency in Education, of which Dr. G. D. Strayer was chairman. The data was secured by questionnaires in the early part of 1919 and the report itself was prepared mainly by Dr. E. S. Evenden. Data were collected from some 382 cities of all sizes, representing all kinds of economic and social conditions in all parts of the country, and from 10 county school systems.

The superintendents filled out questionnaires on the distributions of salaries, on salary schedules, on current salary standards, and on the cost of living in their respective communities. Some 15,000 teachers in these same school systems also filled out questionnaires, giving their personal data on such of the items as affected them. State superintendents of public instruction supplied statements of the general salary situation in their own states. A fourth section of the report deals with teachers' salaries and the cost of living and compares the salaries of teachers with those in similar occupations. The final section deals with the present situation on teachers' salaries, lists elements to be considered in making such schedules, suggests a standard salary schedule for the whole country, gives advice on how to administer such a schedule and concludes with some pertinent observations on the problem of arousing public interest in teachers' salaries. The conclusions are conveniently summarized in point form for each main section. This is a thoroughly workable form but one which does not easily lend itself to a convenient brief digest here.

2. Any school board member, any school executive or any future school executive, who aims to keep abreast of the times educationally, must devote much time and energy to the crucial matter of teachers' salaries. No such person can afford to work without a copy of this report in the school library. Any school executive or any future school executive will have great difficulty in procuring more value for his professional library from $1.50 than by investing that amount in this report.

3. In early 1919 few if any educational problems were more important than finding out the true salary situation in the United States and ways of keeping intact at least skeleton staffs of competent teachers. It was a real emergency in education. Possibly the study of ways of securing more money for schools was of greater importance just then. But even this is debatable since it is practically impossible to secure more money for public schools unless there are at least enough capable teachers to demonstrate clearly to the taxpayers what the latter are expected to raise more money for. The problem was worthy of the conspicuous ability and best efforts of the commission that dealt with it. Indeed it deserved the undivided attention of a much larger group of equal caliber. It fully merited the eight months devoted to it by the man of recognized competence who shaped it for effective presentation to the teaching profession and friends of education throughout the country.

4. The investigation aimed to set standard salary schedules for all grades of teachers for various parts of the country and sizes of cities. It aimed to check up the validity of these standards by securing the facts as to salaries paid teachers, by determining the salaries actually necessary to maintain teachers at reasonable efficiency, by taking into account the trend of the times as seen by the most competent judges (the state superintendents) of conditions over large areas; and by determining what salaries would be necessary to attract capable talent in competition with other occupations. It also aimed to give as accurately as possible a desirable salary schedule applicable to city systems throughout the country and to show city school systems how to enact and operate a desirable salary schedule. All of these aims are markedly achieved in the report.

5. The report will serve admirably as a model for committees investigating salaries in any section of the country, whether in a state or in a smaller area. It points out the different phases to be studied, it shows how the data should be assembled it is full of

valuable suggestions for effective presentation to the public, and its questionnaires may be readily modified for use in local areas.

In addition it has numerous minor conclusions that will suggest elaborations of particular phases of a salary investigation to investigators where such phases constitute particularly acute problems. Thus a normal-school president will derive stimulus from finding on page 98:

> Few states have a salary schedule for their normal schools which will bear comparison with the salaries in the better high schools of the State. As a nation we spend as much per student per year to educate any high school pupil or any boy in the reform school as we do for the training of a teacher in a normal school. We spend nearly twice as much for a pupil in a deaf school, more than twice as much for a college university student, three times as much for a blind student, and nearly four times as much for the training of a feeble minded child.

Anyone interested in combating a plan to employ mainly teachers who are partially supported by their parents, or in securing more men teachers for a given system, will find valuable data and suggestions on pages 76–77. Any superintendent or school board ought to take fresh courage from this on page 132:

> The above study would conclusively indicate that this question of increases to teachers' salaries is largely a matter of local progress, and depends more upon the development of a favorable community attitude or upon the aggressive work of a superintendent or teachers' organization, than upon any economic development of the community. Such a study is evidence of the oft-quoted statement that a community will find the means of supporting schools when convinced that it is a desirable thing to do.

The suggested standards in the table on page 148 as well as the data on actual salaries in 1919, have increased value because of the present nation-wide financial stringency. We are now back to where the 1919 figures represent current conditions more accurately than they did in 1920.

6. The investigation leaves practically untouched the salaries of rural teachers and thus does not directly treat the salary problem as it affects the majority of teachers in the country. But this phase could not have been adequately treated in the time available. Furthermore the variations and allowances for local conditions are much more troublesome to handle for rural teachers. Then too rural salaries have been vitally affected by the draining off of rural teachers to city schools even faster than formerly. Increases in salaries for city teachers have increased rural salaries through increased competition for the better rural teachers. The commission greatly aided rural salaries by its report even if it did not study them directly. And it wisely confined its efforts to what could be adequately studied and reported in time to be of actual value. Aside from this matter of rural salaries, the report leaves little of importance undone. Of course as general conditions change, similar reports will be desirable in the future.

Inquiry No. 1 of the *Know and Help Your Schools* movement conducted by the National Committee for Chamber of Commerce Cooperation with the Public School and by the American City Bureau is such a report and brings the figures up to the summer of 1920.[1]

CARTER ALEXANDER

Wisconsin State Department of Public Instruction

[1] Published by American City Bureau, New York, Chicago, San Francisco, and Toronto.

VAN WAGENEN, MARVIN J. *Historical information and judgment in pupils of elementary schools.* (Teachers College Contributions to Education No. 101) New York: Teachers College, Columbia University, 1919. 74 pp.

Van Wagenen has devised seven scales. Information A, Information B, Thought A, Thought B, Character A, Character B, and Character L. The A and B scales appear not to be exactly equivalent but are intended to be used as duplicate forms. Each of the scales consists of a series of exercises arranged in order of difficulty as determined by the percent of correct responses obtained by submitting them to pupils in grades four to eight. The pupil's score in each case is the number of correct responses.

The author states that "the chief desiderata in instruments for measuring achievement in American history are (1) that the tasks be symptoms of important abilities really desired by the school; (2) that the tasks be not too much disturbed by linguistic difficulties so that ability in history, not in reading or composition, may be chiefly measured; (3) that the measurement of a small group, such as a class of twenty-five or more, be made with sufficient precision; (4) that the tests be capable of extension to alternative forms so as to reduce the harm done by special preparation or coaching for the tests; and (5) that the administration and scoring of the tests be convenient."

The author states that he is leaving the reader to judge for himself concerning the fulfillment of the first two of these requirements. Nowhere in the monograph does he give any account of the construction of the exercises. Usually in making such a group of tests the author finds that it is necessary to reject a number of exercises that have been constructed because they do not yield satisfactory responses when submitted to pupils. Presumably Dr. Van Wagenen must have had this same experience. It is unfortunate that he does not tell us the way in which he arrives at the particular exercises which constitute his tests. One writer[2] has pointed out that a number of the exercises in the thought scales appear to call for factual information. To the extent to which this is true the thought scale is simply another information scale. Thus it is doubtful whether the first two requirements have been satisfied.

The author shows that the other three requirements are met by these scales. The coefficient of reliability, using pupils of the same sex and grade, are 0.71 for information, 0.74 for thought and 0.83 for character. The probable error of these coefficients is given as ± 0.01. The mean square error of a pupil's score is approximately 0.7 for the information and thought scales, and 0.6 for the character scales.

In addition to the derivation of the scale the writer investigated grade, age, and sex differences. The usual individual differences were found in the case of grade and age. In the case of sex differences boys were found distinctly superior to girls.

The general plan of the construction of these scales in that which has been used by Buckingham, Trabue, Woody, and others. The work of Dr. Van Wagenen, therefore, exhibits no originality in this respect. However, if this plan of scale derivation is accepted, as it apparently is, there is no opportunity for originality. As stated above the author fails to tell how the particular exercises of the scales were selected and therefore we cannot know whether there is any originality in this phase of the work or not. So far as one may judge from the published scales this seems doubtful.

W. S. M.

[2] Buckingham, B. R. "A proposed index of efficiency in teaching United States history," *Journal of Educational Research*, 1:161–71, March, 1920.

BETTS, G. H. *How to teach religion: principles and methods.* New York: Abingdon Press, 1919. 223 pp.

There is a growing movement in America toward better methods in religious education. The Protestant churches are waking to the weakness of the policy which relies upon preaching to the neglect of teaching, and undertakes the conversion of adults rather than the Christian nurture and education of children. It has become clear, moreover, that the very magnitude and efficiency of our system of public education may operate, because of its omission of religious teaching, to discredit religion in the minds of the young. If education and religion are not ultimately to be divorced, the churches must assume responsibility for the religious education of children with something like the same clarity of purpose, soundness of method, and efficiency of organization that should characterize the work of the public schools.

The past twenty years have witnessed the introduction of graded lessons into the Sunday schools and the gradual displacement of the old uniform lesson system whereby the same material was taught to children and adults of all ages. The churches are coming to realize, moreover, that the religious education of children involves not merely instruction but activity, and they have begun to match the project methods of the public schools with plans for the comprehensive, motivated education of children in the problems and principles of Christian living. With this has come a rapidly growing demand for week-day religious instruction, and in many communities the public schools are now affording to the churches for this purpose one or more hours a week from their schedule. Many churches now pay some of the teachers in their Sunday schools, and employ directors of religious education, who give their full time to the supervision of the church's plans for the education of its children, young people, and adults. Theological seminaries generally have instituted departments of religious education in which students are trained for the educational aspects of their future work as Christian ministers.

Professor Betts, who was for many years the head of the department of psychology and education at Cornell College, two years ago accepted a professorship of Religious Education in Northwestern University. This book is the first fruitage of his labor in this new field.

All good teaching, Professor Betts holds, rests on a four-fold foundation of principles: those concerned with the aims of teaching, the material or subject-matter of teaching, the organization of the curriculum, and the method or technic of instruction. He discusses "how to teach religion" from each of these four points of view. He defines the aim of religious teaching in terms of knowledge, attitudes, and conduct; and devotes successive chapters to the consideration of what religious knowledge is of most worth, what religious attitudes are to be cultivated, and how religious knowledge and religious attitudes may be used to shape the conduct of daily life. There is a chapter upon the evaluation of materials, biblical and extra-biblical, including materials drawn from nature, history, biography, art, and music. In a chapter on the organization of material there is a brief discussion of the principles underlying the construction of a curriculum of religious education—the problem to which Professor Betts with his graduate students is particularly devoting himself. The last four chapters of the book discuss methods of teaching religion with all the concreteness and insight which have been characteristic of Professor Betts' discussions of like problems of method in the public schools.

Readers of the author's former books, especially "The Mind and Its Education," "The Recitation," and "Social Principles of Education" will expect just the fresh, vigorous, concrete treatment, sound in principle, which is here afforded. Now and then there is evidence that he is at work in a field comparatively new to him. Occasional slips, such as attributing the Epistle of the Hebrews to Paul, show that the author is not as well trained in Biblical literature and theology as in psychology and pedagogy. His criticisms of existing curricula, while for the most part sound, do not bear the marks of actual classroom experience with this material. Professor Betts is not writing for theologians, however, or for lesson-writers, but for the great body of those who teach religion. On the whole his book is the best that we have in its field, and is a very welcome addition to the church's resources as it faces the problem of how to make religion an effective part of the education of children.

L. A. WEIGLE

Yale University

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Walter S. Monroe, University of Illinois, Urbana, Illinois.

Indiana Conference

The conference on educational measurements held annually by Indiana University will take place this year on Friday and Saturday, April 22 and 23. Mr. S. A. Courtis of Detroit will give three addresses before the conference, Friday morning, afternoon, and evening. Dr. Charles H. Judd of the University of Chicago will address the conference on Saturday morning and afternoon. Reports of the discussions will be contributed by members of the Indiana University faculty of the School of Education. This is the oldest and also one of the best known conferences on this topic. The speakers secured for this year insure the usual degree of excellence of the program.

Extension Service of Research Bureaus

Within the last few years a number of university normal schools have established departments, which are given some such titles as "Bureau of Educational Research" or "Bureau of Educational Measurements." In some instances the function of such an organization is considered to be primarily that of research. This research is to be carried on in cooperation with the school men of the state and frequently is based upon problems which are suggested by them. In other instances the major function of the organization of the department has been considered to be that of instructing superintendents, principals, and teachers in the use of educational tests and other scientific methods. This has been done by regular extension classes and by conferences. Recently this department received an account of some of the work which is being carried on this year by the Bureau of Educational Measurements and Standards of the

Kansas State Normal School, Emporia, Kansas. Since this is one of the oldest of such organizations it is interesting to note this statement by the director: "We have given the National Intelligence Test to the children of the elementary schools of Junction City, Fredonia, and Clearwater. Next week we shall give them in Holcomb and Cimarron. Extension classes are being conducted in three centers. We have had more calls for assistance of this type than our force will allow us to handle. Next year we hope to have one man give his full time to field work."

The Kansas Uniform Grading Plan

This plan was indorsed at the City Superintendents' and High School Principals' Sections of the Council of Administration, January 20, 1921. It includes both a definition of school marks and a definition of the distribution of these marks. The plan represents an attempt to make more scientific and accurate the marking of high-school pupils. The plan has been in operation in a number of high schools in Kansas for sometime. The indorsement does not make obligatory the use of the plan but no doubt it will be adopted by a large number of high schools in the state. The plan was formulated by a committee of which W. W. McConell, principal of the high school, Neodesha, Kansas, was chairman.

DEFINITION OF GRADES

GRADE OF A

1. Scholarship—Exceeding expectations of instructor.
2. Initiative—Contributions exceeding the assignment.
3. Attitude—Positive benefit to the class.
4. Cooperation—Forwarding all group activities.
5. Individual Improvement—Actual and noticeable.

GRADE OF B

1. Scholarship—Accurate and complete.
2. Initiative—Stimulating some desirable achievements.
3. Attitude—Proper and beneficial.
4. Cooperation—Effective in group work.
5. Individual Improvement—Showing marks of progress.

GRADE OF C

1. Work in general of medium quality.
2. Work quite strong in one or more items but weak in others.

GRADE OF D

(This grade might be produced by any variety or combination of weaknesses as the definition suggests.)

1. Scholarship—Barely meeting assignments.
2. Initiative—Uncertain, not usually manifest.
3. Attitude—Not objectionable, usually neutral.
4. Cooperation—Not positive nor very effective.
5. Individual Improvement—Slight, not positive.

GRADE OF F

1. This is a failing grade and since it may result from any number of weaknesses is not defined.

DISTRIBUTION OF GRADES

This is a very important factor of this system of grading. Within the class and within the school it is expected that a proper respect be paid to the normal curve of frequency. However, the grades are not to be made primarily for the purpose of fitting the curve. The curve is to be used as a check to see what the "norm" should be under usual conditions and with large enough group. It is expected that a teacher be able to justify his variation in grade distribution from the normal curve of frequency.

NORMAL CURVE OF FREQUENCY

The curve is built upon the following percentages: A—7%, B—24%, C—38%, D—24%, F—7%.

FUNCTION OF THIS SYSTEM

With this code a teacher should grade carefully and be able to defend every grade given. It should be pointed out to the student with a low grade that he is falling short of the standard in some one of these valuable qualities. Teachers have long felt the need of definite points of defense when questioned about a grade but these five desirable qualities need emphasis and should be used in conferences with students. It is essential that students understand this system as thoroughly as do the teachers. This will give students a secure sense of direction in attempting improvement.

In this grading system teachers should bear in mind that grades are given on a comparative basis only; every student's work is to be compared with that of every other student with whom he works, and grades are to be arranged according to his ranking in the class. This will apply to every point under the "definition of grades" as scheduled above. In no case is any grade to be translated to or thought of in terms of percents, but students are to learn that if they make A or B grades, they are considered to belong to approximately the upper one-third of the class; if they are C students, they belong to approximately the middle one-third of the class; and if they rank below C they are among approximately the lower one-third of the class.

WEIGHTED CREDITS

(Suggested by the committee but not recommended unless it received the sanction of the state department)

1. Grade of A for unit subject, 1.2 units.
2. Grade of B for unit subject, 1.1 units.
3. Grade of C for unit subject, 1 unit.
4. Grade of D for unit subject, 0.9 unit.
5. Grade of F for unit subject, no credit.

Training Teachers to Use the New York City Penmanship Scale

Our readers will remember the article by Mr. Nifenecker on "Grade Norms for the New York City Penmanship Scale," which appeared in our December, 1920, number. The Bureau of Reference, Research and Statistics, of which Mr. Nifenecker

is director, has lately succeeded in making up sets of practice material for teachers to use in training their judgment of handwriting by the use of the New York City Penmanship Scale.

This practice material consists of two series of facsimile samples, each series consisting of twenty samples. In addition to this there is a third series of twenty samples which is used as a test series. The two practice series are used in the following manner: The principal distributes a practice series (Series A) to his teachers. The teachers then rate the samples of this series against the scale, taking account of form, movement, and spacing. After the teachers have done this their results are checked by the principal against the known values of the samples which have been furnished to him. A conference with each teacher ensues during which the greatest variations from the standards are taken up. Again the teacher, using the practice series, will rate the samples against the scale. This procedure is carried on until the teacher's variation from the established values of the samples has been materially reduced. As a final check on the accuracy of the teacher's judgment the test series of samples is used.

As has been indicated above, an alternative practice series (Series B) is also provided. It is to be used when familiarity with the A series has made its further use undesirable.

This follow-up of the New York City Penmanship Scale on the part of the Bureau of Reference, Research and Statistics constitutes a fine example of carrying through a research project to the point not only of installation but also of supervision.

Correlation of Certain Intelligence Tests for the Lower Grades

In these columns of a recent issue of the Journal appeared a discussion of the Illinois General Intelligence Scale based upon data furnished by Superintendent L. W. Keeler of Michigan City, Indiana. From the same source the Bureau of Educational Research has also received scores of a number of lower grade pupils on several tests. About 20 pupils in the first grade, an equal number in the third grade, and 35 in the second grade were given the Stanford Revision of the Binet-Simon Tests, the Dearborn Group Tests of Intelligence, and the Kingsbury Primary Group Intelligence Scale. For purposes of correlation the mental ages in months on the first mentioned, the average point scores on the second, the point scores on the last, and the average of these two latter scores were used. Throwing the scores for the pupils of the three grades together, the following coefficients were obtained:

Stanford with Dearborn	0.87
Stanford with Kingsbury	0.73
Stanford with Average	0.87
Dearborn with Kingsbury	0.75
Dearborn with Average	0.92
Kingsbury with Average	0.89

Assuming that the Stanford Revision is a standard for group intelligence tests, these figures imply that both the other tests are rather accurate, the Dearborn being so to a high degree and as good as the average of the two.

The correlation coefficients for the three grades separately were also computed. They were:

	Grades		
	I	II	III
Stanford with Dearborn	0.66	0.78	0.60
Stanford with Kingsbury	0.32	0.53	0.60
Stanford with Average	0.61	0.71	0.66
Dearborn with Kingsbury	0.53	0.54	0.47

These coefficients indicate that the superiority of the Dearborn Test is with the younger pupils, but that later it possesses no advantage over the Kingsbury.

However, the matter may be looked at from another viewpoint. The following table gives the means, medians, standard deviations and coefficients of variation (Pearson's formula).

	Grades			Grades Continued
	I	II	III	
	Medians			
Dearborn	23	44	59	43
Kingsbury	21	30	53	30
	Means (Averages)			
Dearborn	27	44	59	43
Kingsbury	22	32	59	36
	Standard Deviations			
Dearborn	9.7	12.8	10.0	15.9
Kingsbury	6.2	16.3	21.3	20.5
	Coefficients of Variation			
Dearborn	36	29	17	37
Kingsbury	28	51	36	57

Since the two tests are both scored on a scale of 100 points and the distributions obtained in the two are located not far from the same portion of the scale, standard deviations and coefficients of variation seem to indicate that in grades II and III, the Dearborn Tests are suitable for testing a group of greater range of ability than is the Kingsbury Scale, but that the latter is scored in smaller units and hence discriminates more finely. In grade I the reverse seems true. The increases from grade to grade, which, as judged by the Dearborn, are almost equal, while by the Kingsbury that from grade I to II is less than one-half of that from II to III, also point to this conclusion.

There is another possible interpretation of the data given above, however. Other things being equal, the greater the accuracy of a test the less its variation. Though our data are not sufficient to say that all other things are equal, yet since both tests cover a range of 100 points and the averages of the distributions are fairly close to each other, these facts furnish some evidence that the Dearborn is the more accurate in the two upper grades, and the Kingsbury in grade I.

On the whole, the evidence obtained from the coefficients, both of correlation and of variation, and from the increases in score from grade to grade seems to point to the

fact that the Dearborn is somewhat superior to the Kingsbury for the three grades mentioned. Whether or not this superiority is great enough to justify the at least double expenditure of time required in its administration is a question the writer prefers not to attempt to answer.

C. W. ODELL

University of Illinois

A Method of Tabulating Scores in Group Testing

During the early fall of 1920 the seventh and eighth grades of two Boston schools were given the Dearborn Group Tests of Intelligence. After the tests had been corrected, it was decided that it was unwise to report the results to teachers and administrators in the form of intelligence quotients. It was believed that the chance of misinterpreting such returns by teachers and parents was great enough to warrant seeking some new method of making a report to the school.

After trying various ways of treating the distributions, it was decided that the method described in this communication met all the necessary requirements and made it possible to locate each pupil within certain ranges, which were near enough to actual conditions for all practical purposes.

After the papers were corrected a distribution was made of all scores of each grade tested regardless of age. The median, and the first, and third quartiles of the distribution for each grade were then determined. Following this, a distribution of scores on a basis of age regardless of the grade was made. That is, all pupils were divided into age groups of six months each, 10 years to 10 years, 5 months; 10 years, 6 months to 10 years, 11 months; 11 years to 11 years, 5 months, etc. Then a distribution of the scores for each of these age groups was made and the median, and first, and third quartiles were found. Then a third distribution was made in which the pupils were grouped according to ages. The median age, and the first, and third quartiles were found.

It was now possible to place each pupil in his proper quarter according to each of these distributions and thus to obtain his grade-score, his age-score, and his chronological age (in terms of quarters). Table I and Table II show the way each class was reported to the school. The first quarter shows a high score or high age and the fourth quarter shows low score or low age. Under pupils' names letters have been used instead of names to indicate pupils. This report shows that most of the pupils fell in the first or second quarter of the distribution in both grade score and age score with only one (Z) falling in the fourth quarter. Sixteen are young for their grade as shown by the fact that they are in the fourth quarter and only four are old for their grade as shown by the fact that they are in the first quarter. In other words, this is a superior class.

TABLE I. SHOWING SEX, AGE, SCORE, AND QUARTILE STANDING OF 32 EIGHTH-GRADE PUPILS, ROOM 346

Sex	Pupil's Name	Age	Pupil's Score	Quartile Standing		
				Grade Score	Age Score	Age
B	A	13-1	68	1	1	3
G	B	12-5	66	1	1	4
B	C	12-3	64	1	1	4
G	D	12-10	71	1	1	3
G	E	13-1	61	1	1	3
G	F	13-5	72	1	1	2
B	G	13-3	68	1	1	2
B	H	13-3	72	1	1	2
G	I	12-4	65	1	1	4
B	J	13-3	61	1	1	2
G	K	12	73	1	1	4
B	L	12	71	1	1	4
B	M	12-4	57	2	2	4
B	N	12	58	1	2	4
B	O	12-6	78	1	1	4
G	P	12-10	81	1	1	3
G	Q	12-3	69	1	1	4
G	R	11-10	87	1	1	4
B	S	12	59	1	1	4
B	T	13-3	59	1	2	2
G	U	11-11	45	3	3	4
G	V	12-7	69	1	1	3
G	W	13-10	53	2	2	1
G	X	11-11	62	1	1	4
B	Y	12-6	69	1	1	4
B	Z	15-6	30	4	4	1
G	AA	14-2	61	1	1	1
G	BB	12-3	55	2	2	4
B	CC	12-3	83	1	1	4
G	DD	12-11	79	1	1	3
G	EE	13-5	51	2	2	2
B	FF	13-8	78	1	1	1

TABLE II. SHOWING SEX, AGE, SCORE, AND QUARTILE STANDING OF 40 EIGHTH-GRADE PUPILS, ROOM 347

Sex	Pupil's Name	Age	Pupil's Score	Quartile Standing		
				Grade Score	Age Score	Age
B	A	14–2	69	1	1	1
B	B	13–10	71	1	1	1
B	C	13	48	4	3	3
B	D	14–2	69	1	1	1
B	E	15–1	40	4	3	1
B	F	14–3	75	1	1	1
B	G	13–5	45	4	3	2
B	H	13–9	65	1	1	1
B	I	14–5	29	4	4	1
B	J	13–8	77	1	1	1
B	K	14–2	64	1	1	1
B	L	13	46	3	3	3
B	M	13–5	57	2	2	2
B	N	13–10	25	4	4	1
B	O	12–11	35	4	4	3
B	P	13–10	37	4	4	1
B	Q	12–10	31	4	4	3
B	R	14	48	3	2	1
B	S	14–11	50	2	2	1
B	T	14–2	33	4	4	1
B	U	18–3	16	4	.....	1
B	V	13–5	41	4	4	2
B	W	13–11	20	4	4	1
B	X	15–9	47	3	2	1
B	Y	15	55	2	1	1
B	Z	14–2	56	2	2	1
B	AA	13–10	63	1	1	1
B	BB	17	31	4	.....	1
B	CC	14–1	41	4	3	1
B	DD	15–1	41	4	3	1
B	EE	13–2	41	4	4	2
B	FF	14–7	52	2	2	1
B	GG	13–4	39	4	4	2
B	HH	14–4	53	2	2	1
B	II	14–7	45	3	3	1
B	JJ	14–5	49	3	2	1
B	KK	12–5	41	4	3	4
B	LL	13–6	41	4	3	2
B	MM	15–9	49	3	1	1
B	NN	13–6	42	4	3	2

In Table II one-half of the pupils fall in the third or fourth quarter of the distribution in both grade-score and age-score, while thirteen attain a score which places them in the first or second quarter. One pupil is young for his grade as shown by the fact that he is in the fourth quarter and 28 pupils are old for their grade as shown by the fact that they are in the first quarter of the distribution. In other words, these pupils are mentally inferior to the children in room 346. If, however, the tests are telling the truth, room 347 is poorly graded as shown by the range of scores from 16 to 77.

A. W. KALLOM

Department of Educational Investigation and Measurement, Boston, Massachusetts

Norms for the Otis Group Intelligence Scale

The norms for the Otis Group Intelligence Scale presented by Dr. Colvin in the January number of the JOURNAL OF EDUCATIONAL RESEARCH cause me some concern. As he mentions, they differ from the norms published by Dr. Otis, but what is of more importance to us locally is that they differ still more from the norms we have derived from local data. The following table shows the differences:

Age	Otis Norms	Colvin Norms	Cincinnati Norms
8	40	58	30(4)
8½	46	70½	36(9)
9	52	67½	42(22)
9½	58	78½	48(66)
10	64	80½	54(67)
10½	70	90	60(91)
11	76	94	66(105)
11½	82	109	72(111)
12	88	118	78(101)
12½	94	119½	84(91)
13	100	130	90(101)
13½	106	133	96(112)
14	112	135½	102(85)
14½	117	134½	108(49)
15	121	136	114(40)

The figures in parenthesis indicate the number of pupils at each age taking the test. Two schools were selected for the testing, upon which these norms are based, by conference with the supervisors. One school was considered a little below average and the other a little above. A combination of the records from both schools was thought to be representative of Cincinnati. Our norms were derived before the final norms of Dr. Otis were published. The work was carefully done and the results have agreed very satisfactorily with Stanford results and with grade placement.

The use of either the Otis or Colvin norms would give us absurd results. We are not inclined to believe that Cincinnati school children are so much lower than other children. The only possible explanation which occurs to us is in the method of giving and scoring. It may not be out of place therefore to mention just how we do our testing.

All of the tests were given by myself and one assistant. The test papers were also scored by us. In the giving, we have been very rigid in following faithfully the printed directions, never varying by way of additional explanation in the slightest degree. We have been equally exact about the scoring, although there is less tendency to vary procedure here than in the giving. Much of our work has been checked, so that we are sure no accidental errors have crept in.

My experience in training teachers to give standard educational tests is that there is an almost irresistible tendency to help a little more than the directions permit. A teacher may read over the directions and believe that a slight change in wording would make the meaning clearer, or think that a certain amount of coaching in taking such a type of test is legitimate. Also, during the test, she has the feeling that the directions have not been sufficiently clear and she adds a word. All of this is possible and likely to happen in more cases than we like to think. The result, of necessity, is to raise the pupils' scores and, if these results are used to make up norms, to raise the norms.

If I am right in this explanation, it is evident that a single set of norms cannot be of universal application. Each locality will have to establish its own norms under the conditions under which it proposes to administer the tests. Granting that our testing procedure could be better standardized, that alone would not help materially unless those who administer the tests could be made to adhere more carefully to the same procedure.

With the rapid increase in testing, are not some very bad and misleading results to be expected unless there is much more care taken to understand the conditions under which the testing is done?

WARREN W. COXE

Assistant Director, Vocation Bureau, Board of Education, Cincinnati, Ohio

The Van Wagenen Reading Scales in History, General Science, and English Literature

Although there are now available several reading scales and several standardized or diagnostic tests for measuring reading ability, little attention so far seems to have been given to the measurement of the ability to read and comprehend the content of any particular one of the school subjects. The present series of reading scales is an attempt to meet just this need. With them the teacher of a special subject, such as history, general science, or English literature, will be able to determine how well the individual students of the group can grasp the content of the material assigned for reading and study. Not a few of the failures to do successful work are found to be due to inability to read and comprehend the ideas in the lessons assigned. With the use of such reading scales it may also be possible to determine more adequately than can be done at the present time the abilities of students to take up certain subjects of study rather than others.

Another and equally important value of the scales lies in the possibility of determining the degree of difficulty of reading material that may safely be placed in the hands of students of various degrees of capacity, ability, and maturity with an expectation that the content will be comprehended. Few indeed are the textbooks in use at the present time that are written consistently within the range of comprehension of the students for whom they are designed. Nor could the situation very well be otherwise in the absence of more definite models for each particular stage of achievement.

No less important but less evident is the value of such scales in making possible (1) relatively accurate comparisons of ability in reading in various subjects; (2) relatively accurate measurements of the effects of increasing the student's range of information in a subject upon his ability to read intelligently in that subject; and (3) relatively accurate measurements of the effect of different methods of study upon the ability to read independently in a certain subject.

The unit of measurement in these scales is the deciquartile or one-tenth quartile. The quartile is the amount by which two exercises must differ in difficulty in order that one shall be done with the same degree of correctness by the mid-pupil in a normally distributed group as the second is done by the pupil who stands at the top of the lower quarter of the group or at the top of the lower three-quarters of the group. The deciquartile, or one-tenth quartile, thus represents one-tenth of the difference in reading ability between the median pupil and either the twenty-five or seventy-five percentile pupil in a normally distributed group in reading ability in a special subject like history. Inasmuch as the zero point in reading ability has not yet been adequately determined, an arbitrary zero point has been selected which lies eighty deciquartiles or eight quartiles below the ability of the median first-year high-school student at the end of the year. Although it will not yet be possible to say that one student has one and one-half times as much reading ability in history as another student has, or that he can read one and one-half times as well in history as in general science, nevertheless one can say that one student or one group has made one and one-half times as much gain as another student or group of students. An illustration would be the case of one group making a gain from 72 deciquartiles up to 80 deciquartiles, while a second group made a gain from 72 deciquartiles up to 84 deciquartiles during the same period.

The exercises in the several reading scales were scaled in terms of the quartile by practically the same method as used by the author in making his American history tests. The percents of students of each group taking the original tests—approximately 750 in grades eight to twelve—who did each exercise correctly were found. With the use of the table devised by W. F. Sheppard, these percents were turned into quartile values. The quartile values for the different grades were then referred to the median of the first-year class, which as stated above, was placed at 80 deciquartiles above an arbitrary zero point. The only difference in method lay in the determination of the distances between the group medians. In this study the differences were based on the averages of distances for each exercise instead of on the overlapping of the groups upon one another's medians, as the former method seemed to give a more adequate measure of the grade differences for the reading exercises.

In connection with the derivation of these scales two other facts should be noted. The scales are purely reading scales. No writing other than the making of check marks is necessary. Each exercise, of which there are fifteen in each scale, consists of a

carefully selected paragraph with a list of from four to six statements below it. These statements—which for the eighth grade and high-school student are of approximately equal difficulty within any one set—either contain an idea that is expressed in the paragraph in different words or that can be derived from a reading of the paragraph, or else contain an idea that might be but is not in the paragraph or derivable from it. The student has merely to check those statements which contain an idea that is to be found in the paragraph or that can be inferred from it. All the teacher needs to do in correcting the test papers is simply to note from a key the number of errors made in each set and to write this number of errors in a special form provided for it at the top of the front page.

Along with the ease with which the exercises may be scored is the ease with which the student's score in terms of a scale value may be calculated. As pointed out by Dr. Truman Lee Kelley, the most reliable scale score for an individual is that score where the amount of error and the amount of correctness are the same. Although the calculation of this point is somewhat involved, the score may be arrived at in a very few seconds with the use of tables recently worked out for this purpose. The data in the original tables have been adapted for use with these scales so that, having the number of errors tabulated in the form provided on the front page of the scale, the student's score may be readily determined merely by consulting two short keys and making one simple addition and one simple subtraction.

From having the scores thus expressed in terms of a scale unit, several advantages accrue. Comparisons of achievement in terms of definite amounts may be made, a process that cannot be done with accuracy with the use of ordinary standard tests made up of unscaled data. The scores made by students on one scale are of the same magnitude and value as those made on another similar scale, even though the starting and ending values of the exercises in the two scales differ. This is a feature that is possible in the usual scale only when the difficulties of the tasks in the duplicate tests are identical—a feature that is not only difficult to attain but also one seldom achieved. From the foregoing statements it is apparent that this series of scales possesses features that should characterize every educational test or scale of this type; namely, there are duplicate scales for each of the subjects so that accurate comparisons of achievement may be made; also additional duplicate scales can be developed with very much less work than the two original scales in each subject required.

M. J. VAN WAGENEN

University of Minnesota

National Association of Directors of Educational Research

(E. J. Ashbaugh, *Secretary and Editor*)

REPORT ON THE SEVENTH ANNUAL MEETING OF THE NATIONAL ASSOCIATION OF DIRECTORS OF EDUCATIONAL RESEARCH

Six years ago at Cincinnati during the meeting of the Department of Superintendence a little group of men, engaged in the relatively new field of educational research in public schools, met for a round table discussion of the problems which were facing them. Believing that an organization of the men of common interests in this field would be both helpful to the men themselves and stimulating to the country at large, they formed a preliminary organization and named committees to draw up a constitution and launch the movement.

At Detroit the following year this initial group with a few others came together for a dinner and an evening of discussion. The next year (1917) at Kansas City a half-day session public program was offered and evoked enough interest to warrant the preparation of a yearbook, which was published by the National Society for the Study of Education as its Seventeenth Yearbook, Part 2. The reception of this yearbook on the part of the public justified the effort of the association in that direction, and the growing numbers attending the public programs have continued to testify to the quality of the material presented upon them.

At the Atlantic City meeting three years ago another public program was offered. Apparently no new developments came about at this meeting except that a new secretary was elected—George Melcher having served from the beginning and having asked to be relieved—and the first provision for regular annual dues was made. The fact is that at this meeting the members were so much occupied in congratulating themselves on the yearbook (which had just appeared) that new ideas were not forthcoming.

Two years ago the members of the association, feeling that they were unable at the public meetings to get all of the intimate critical give-and-take upon their problems which they had originally expected, provided for a one-day meeting open to members only. The first meeting of this sort was held at Cleveland last year.

The second annual closed meeting was held Saturday, February 26, 1921 in Haddon Hall Hotel. About thirty were present, and the meeting was thoroughly informal.

Dr. McCall reported on the development of scale construction with special reference to the new Thorndike-McCall Reading Scales. He called attention to the fact that the maker of an educational measurement scale might employ one of the following five methods: (1) percentile; (2) variability of adult judgment; (3) variability of adult performance; (4) age scale method; (5) variability of pupil performance. He presented the objections to each of these methods as they occurred to him in the making of his reading scales for elementary-school pupils.

The objections are herewith listed practically without discussion, being numbered in the same order as that in which the methods are named:

(1) Objections—(a) varying zero point and (b) unequal units.

(2) and (3) Inappropriate for an elementary reading scale.
(4) (a) Inequality of units; (b) educational growth not equal for different years; result sometimes negative.
(5) (a) Usually involves an attempt to locate zero point, but no agreement or means of reaching agreement in finding it or determining what zero ability is; (b) yields relatively ephemeral results; and (c) in order to produce an even step scale, a large proportion of material must be thrown away.

Therefore, there is need for a method that will have units comparable within the scale itself; units comparable with units of other scales; and a constant reference point. His solution of these problems is given in the latest issue of the *Teachers' College Record*.

Miss Race, psychological examiner for the Kansas City, Missouri, public schools, next presented some experimental data on intelligence scores and silent reading scores.

Saturday afternoon Dr. Rugg reported for the subcommittee on standardization of statistical terms, and suggested that this association should not act until the committee of the National Research Council which is working on the same problem shall have made its report.

Dr. Ashbaugh reported for the subcommittee on child accounting; and Professor Courtis reported for the general committee on standardization. Courtis showed that the returns from his questionnaire were distinguished rather for their differences than for their unanimity of opinion, though the membership tended to divide into two fairly definite groups. He recommended that a decision on policy be based upon the discussion by the members present rather than upon the majority vote he had obtained on the propositions submitted.

A committee, of which Courtis was chairman, was therefore appointed to report to the association on Thursday.

Two actions were taken, however, of which our members should take special note:

1. Motion carried that the reliability of a test shall be understood to be the degree of correspondence between successive scores of the same individuals on the same test.

2. Motion carried that in reference to correlation, members of the association use the term "Product Moment Method" instead of the term "Pearson Method."

On Thursday the open meeting of the association was held. An account of it is given below. Because large numbers of the superintendents as well as some of our members had to leave Atlantic City during the afternoon or early evening, it was decided to change our annual dinner to a noon luncheon. When President Haggerty asked for roll call, three honorary members (J. McKeen Cattell, E. L. Thorndike, and J. M. Rice) and thirty active members responded as follows:

Active Members

S. A. Courtis	Arthur S. Otis	Ira J. Bright
B. T. Baldwin	J. Freeman Guy	Elmer K. Sexton
G. M. Whipple	R. L. Morton	H. A. Greene
Laura Zirbes	H. C. Daley	Thomas J. Kirby
Edna Gordon	J. Cayce Morrison	Charles Fordyce
P. C. Packer	W. W. Coxe	Emma M. Brown
M. E. Haggerty	M. R. Trabue	Harriet M. Barthelmess
E. J. Ashbaugh	George Melcher	B. R. Buckingham
A. B. Wight	Carter Alexander	Henrietta V. Race
L. A. Williams	W. J. Osburn	Harold O. Rugg

Due to the limited time at our disposal, necessary items of business were cared for at intervals during the progress of the luncheon. The president first announced the names of the three men who were elected to honorary membership this year. These were: Dr. Lotus D. Coffman, Dr. Clark Wissler, and Dr. J. McKeen Cattell. Dr. Cattell was the only one of these three men who were able to be with us at the luncheon.

The secretary also announced that during the past year there had been twenty-three members added to the active roll. The complete membership list will appear in the next issue of the Journal.

A communication from Dr. L. P. Ayres, a charter member of our association, containing his resignation from membership was read. The association voted to accept his resignation with regret, and elected him immediately to honorary membership.

The committee appointed on Saturday to revise the membership article of the constitution reported, and after a slight modification had been made, the report was accepted. The new article will appear in this department in the next issue.

The committee on standardization rendered a report which the association adopted and ordered printed in the Journal at an early date. Action was also taken requiring that the report on child accounting which was read on Saturday should be sent out to the members for vote, and the committee on standardization was given discretionary power to print this report also.

The president's annual address confined itself largely to the progress in intelligence testing. It was received with so much favor that the association requested that it be printed in the Journal for the benefit of our entire membership.

The following were elected for the ensuing year:

President—Dr. H. O. Rugg, Educational Psychologist, Lincoln School, Teachers College, Columbia University, New York City.

Vice-President—Dr. Virgil E. Dickson, Director of Bureau of Educational Research, Oakland and Berkeley, California.

Secretary-Treasurer—Dr. E. J. Ashbaugh, Director of Bureau of Educational Service, State University of Iowa, Iowa City, Iowa.

(The following write-up of the Thursday program was contributed by our accommodating editor, B. R. Buckingham.—E. J. A.)

After dragging through the intervening days and conscientiously exposing ourselves to the warmed-over program materials offered at other meetings, it was a relief to attend our Thursday meeting. Nor were we alone in our notion. Although President Haggery had scheduled the meeting for the unusually early hour of nine o'clock, the hall was fairly well filled when the first paper was announced, and by ten o'clock we were playing to standing room only.

The program moved forward with clock-like precision. Every person who was scheduled for an address was present. Each paper was given precisely in the order of its appearance on the printed program.

Mr. Buckingham's "One Hundred Percent Promotion" showed how actual promotion rates of 95 percent were secured in two cities for three successive semesters without detriment to grade standards. Mr. Fordyce presented a capital paper on the classification of children through the use of intelligence tests. Mr. Whipple let us into some of the mysteries in the derivation of the National Intelligence Tests. His observations on

the value of these tests and the uses that may be made of them were well received. Mr. Bright showed some of the uses he has been making of intelligence examinations with high-school freshmen. His results were related to the success of pupils in several school subjects. Mr. Thorndike distributed copies of his Intelligence Examination and gave a highly suggestive talk on the testing of high-school graduates, using this examination as his text. This ended the morning program. President Haggerty had planned to close the session at 11:45 and the smooth-running machine which he had set up came to a halt exactly as scheduled.

By telescoping business with eating we were able to accomplish a great deal of both. The afternoon program was to begin at two o'clock in another building. The luncheon, however, lasted until nearly two-thirty. The president, not being endowed with the property of ubiquity, dispatched Mr. Fordyce to open the meeting and introduce Mr. Gray—W. S. of Chicago—to the expectant audience. Any expectations which the audience may have entertained were amply justified. For concrete, helpful data such as could be used in the classroom, Mr. Gray's paper on "Diagnostic and Remedial Steps in Reading" was only rivaled by Mr. Horn's on "A Constructive Program in the Development of Ability to Read Complicated Materials." Indeed, so rich were these papers in what Horn calls "factual" material that nothing short of the careful reading of their printed reports will be fully satisfactory. It is a tribute to these papers that the audience was frantically taking notes while the speakers were delivering them.

Mr. Rugg gave a paper on "The Curriculum in Economic and Social Studies." In this paper he maintained that teachers get little further than the textbook in securing materials for instruction, that textbooks were largely written by college men, and that the real sources from which the curriculum should be derived are the needs and uses of society. Mr. Alexander gave a highly interesting sketch of some of the factors influencing the successful presentation to the public of educational measurements. The "fact basis" of this discussion was in part the material which the State Department of Education of Wisconsin has gathered concerning publicity campaigns. He gave many important suggestions and cautions for publicity work.

Secretary Ashbaugh at this point tore himself away from the scribbling on which he had been mysteriously engaged and presented a paper on "The Measurement of Language." He reviewed the development of the measurement of composition and language from the very general scales of the Hillegas family through the more analytic instruments of Ballou, Trabue, and Green to those of Starch, Charters, and Kirby. He asserted that this analysis should go further and that instruments which get at specific details are the only ones which can be relied upon to reveal rather than conceal the defects of pupils in language.

The valedictory was pronounced by Mr. Courtis who presented "Research as a Distinct Function of Supervision." He showed clearly the truly remarkable type of organization which has been developed at Detroit. It is quite impossible within the limits of this sketch to say anything adequate concerning it. We can only hope that Mr. Courtis will give himself full scope in telling us about it when he submits his program article for publication.

We cannot conclude our account of the meetings of the association without an expression of appreciation to President Haggerty. The members of the Research Association and the public feel that an exceptionally scholarly and helpful program was presented as the result of his efforts.

JOURNAL *of* EDUCATIONAL RESEARCH

VOLUME III MAY, 1921 NUMBER 5

EDUCATIONAL AIMS, IDEALS, AND ACTIVITIES

W. W. CHARTERS
Carnegie Institute of Technology

Aims precede the curriculum. Alike in the organized writings of educators and in the changing conceptions of unorganized educational opinion, fundamental changes in the curriculum occur only after fundamental modifications of the aims of education have been discovered. Among the writers, Plato described the qualities of the perfect warrior guardian of the state before determining upon his education; Comenius presents the aims of virtue, piety, and learning before prescribing the curriculums of his systems of education; and Rousseau justifies his course of study by appeal to certain pre-formed assumptions concerning "the nature" of the child. Likewise, in the fluid movement of thought of the present generation when the curriculum of our times is in the making, we find that only after such aims as social efficiency have been subjected to prolonged discussion, do changes in the curriculum begin to occur.

This precedence of aims over subject matter is a necessity for purposes of selection and use. Because the information which youth inherits is beyond his powers of complete assimilation, selection must be made and a line must be drawn between the more useful and the less useful. And since selection is based upon value, educational aims must first be formulated. Moreover, since the functions of subject matter must be taught (for to present facts without an acquaintance with the uses to which they are to be put is futile) in the last analysis the use of any individual fact must be derived from the central aims of the whole system of education. So, for both selection and use the aims of education must precede the construction of the curriculum.

(2) and (3) Inappropriate for an elementary reading scale.

(4) (a) Inequality of units; (b) educational growth not equal for different years; result sometimes negative.

(5) (a) Usually involves an attempt to locate zero point, but no agreement or means of reaching agreement in finding it or determining what zero ability is; (b) yields relatively ephemeral results; and (c) in order to produce an even step scale, a large proportion of material must be thrown away.

Therefore, there is need for a method that will have units comparable within the scale itself, units comparable with units of other scales; and a constant reference point. His solution of these problems is given in the latest issue of the *Teachers' College Record*.

Miss Race, psychological examiner for the Kansas City, Missouri, public schools, next presented some experimental data on intelligence scores and silent reading scores.

Saturday afternoon Dr. Rugg reported for the subcommittee on standardization of statistical terms, and suggested that this association should not act until the committee of the National Research Council which is working on the same problem shall have made its report.

Dr. Ashbaugh reported for the subcommittee on child accounting; and Professor Courtis reported for the general committee on standardization. Courtis showed that the returns from his questionnaire were distinguished rather for their differences than for their unanimity of opinion, though the membership tended to divide into two fairly definite groups. He recommended that a decision on policy be based upon the discussion by the members present rather than upon the majority vote he had obtained on the propositions submitted.

A committee, of which Courtis was chairman, was therefore appointed to report to the association on Thursday.

Two actions were taken, however, of which our members should take special note:

1. Motion carried that the reliability of a test shall be understood to be the degree of correspondence between successive scores of the same individuals on the same test.

2. Motion carried that in reference to correlation, members of the association use the term "Product Moment Method" instead of the term "Pearson Method."

On Thursday the open meeting of the association was held. An account of it is given below. Because large numbers of the superintendents as well as some of our members had to leave Atlantic City during the afternoon or early evening, it was decided to change our annual dinner to a noon luncheon. When President Haggerty asked for roll call, three honorary members (J. McKeen Cattell, E. L. Thorndike, and J. M. Rice) and thirty active members responded as follows:

Active Members

S. A. Courtis	Arthur S. Otis	Ira J. Bright
B. T. Baldwin	J. Freeman Guy	Elmer K. Sexton
G. M. Whipple	R. L. Morton	H. A. Greene
Laura Zirbes	H. C. Daley	Thomas J. Kirby
Edna Gordon	J. Cayce Morrison	Charles Fordyce
P. C. Packer	W. W. Coxe	Emma M. Brown
M. E. Haggerty	M. R. Trabue	Harriet M. Barthelmess
E. J. Ashbaugh	George Melcher	B. R. Buckingham
A. B. Wight	Carter Alexander	Henrietta V. Race
L. A. Williams	W. J. Osburn	Harold O. Rugg

Due to the limited time at our disposal, necessary items of business were cared for at intervals during the progress of the luncheon. The president first announced the names of the three men who were elected to honorary membership this year. These were: Dr. Lotus D. Coffman, Dr. Clark Wissler, and Dr. J. McKeen Cattell. Dr. Cattell was the only one of these three men who were able to be with us at the luncheon.

The secretary also announced that during the past year there had been twenty-three members added to the active roll. The complete membership list will appear in the next issue of the Journal.

A communication from Dr. L. P. Ayres, a charter member of our association, containing his resignation from membership was read. The association voted to accept his resignation with regret, and elected him immediately to honorary membership.

The committee appointed on Saturday to revise the membership article of the constitution reported, and after a slight modification had been made, the report was accepted. The new article will appear in this department in the next issue.

The committee on standardization rendered a report which the association adopted and ordered printed in the Journal at an early date. Action was also taken requiring that the report on child accounting which was read on Saturday should be sent out to the members for vote, and the committee on standardization was given discretionary power to print this report also.

The president's annual address confined itself largely to the progress in intelligence testing. It was received with so much favor that the association requested that it be printed in the Journal for the benefit of our entire membership.

The following were elected for the ensuing year:

President—Dr. H. O. Rugg, Educational Psychologist, Lincoln School, Teachers College, Columbia University, New York City.

Vice-President—Dr. Virgil E. Dickson, Director of Bureau of Educational Research, Oakland and Berkeley, California.

Secretary-Treasurer—Dr. E. J. Ashbaugh, Director of Bureau of Educational Service, State University of Iowa, Iowa City, Iowa.

(The following write-up of the Thursday program was contributed by our accommodating editor, B. R. Buckingham.—E. J. A.)

After dragging through the intervening days and conscientiously exposing ourselves to the warmed-over program materials offered at other meetings, it was a relief to attend our Thursday meeting. Nor were we alone in our notion. Although President Haggery had scheduled the meeting for the unusually early hour of nine o'clock, the hall was fairly well filled when the first paper was announced, and by ten o'clock we were playing to standing room only.

The program moved forward with clock-like precision. Every person who was scheduled for an address was present. Each paper was given precisely in the order of its appearance on the printed program.

Mr. Buckingham's "One Hundred Percent Promotion" showed how actual promotion rates of 95 percent were secured in two cities for three successive semesters without detriment to grade standards. Mr. Fordyce presented a capital paper on the classification of children through the use of intelligence tests. Mr. Whipple let us into some of the mysteries in the derivation of the National Intelligence Tests. His observations on

the value of these tests and the uses that may be made of them were well received. Mr. Bright showed some of the uses he has been making of intelligence examinations with high-school freshmen. His results were related to the success of pupils in several school subjects. Mr. Thorndike distributed copies of his Intelligence Examination and gave a highly suggestive talk on the testing of high-school graduates, using this examination as his text. This ended the morning program. President Haggerty had planned to close the session at 11:45 and the smooth-running machine which he had set up came to a halt exactly as scheduled.

By telescoping business with eating we were able to accomplish a great deal of both. The afternoon program was to begin at two o'clock in another building. The luncheon, however, lasted until nearly two-thirty. The president, not being endowed with the property of ubiquity, dispatched Mr. Fordyce to open the meeting and introduce Mr. Gray—W. S. of Chicago—to the expectant audience. Any expectations which the audience may have entertained were amply justified. For concrete, helpful data such as could be used in the classroom, Mr. Gray's paper on "Diagnostic and Remedial Steps in Reading" was only rivaled by Mr. Horn's on "A Constructive Program in the Development of Ability to Read Complicated Materials." Indeed, so rich were these papers in what Horn calls "factual" material that nothing short of the careful reading of their printed reports will be fully satisfactory. It is a tribute to these papers that the audience was frantically taking notes while the speakers were delivering them.

Mr. Rugg gave a paper on "The Curriculum in Economic and Social Studies." In this paper he maintained that teachers get little further than the textbook in securing materials for instruction, that textbooks were largely written by college men, and that the real sources from which the curriculum should be derived are the needs and uses of society. Mr. Alexander gave a highly interesting sketch of some of the factors influencing the successful presentation to the public of educational measurements. The "fact basis" of this discussion was in part the material which the State Department of Education of Wisconsin has gathered concerning publicity campaigns. He gave many important suggestions and cautions for publicity work.

Secretary Ashbaugh at this point tore himself away from the scribbling on which he had been mysteriously engaged and presented a paper on "The Measurement of Language." He reviewed the development of the measurement of composition and language from the very general scales of the Hillegas family through the more analytic instruments of Ballou, Trabue, and Green to those of Starch, Charters, and Kirby. He asserted that this analysis should go further and that instruments which get at specific details are the only ones which can be relied upon to reveal rather than conceal the defects of pupils in language.

The valedictory was pronounced by Mr. Courtis who presented "Research as a Distinct Function of Supervision." He showed clearly the truly remarkable type of organization which has been developed at Detroit. It is quite impossible within the limits of this sketch to say anything adequate concerning it. We can only hope that Mr. Courtis will give himself full scope in telling us about it when he submits his program article for publication.

We cannot conclude our account of the meetings of the association without an expression of appreciation to President Haggerty. The members of the Research Association and the public feel that an exceptionally scholarly and helpful program was presented as the result of his efforts.